The Dissertation Committee for John Michael Leavitt Certifies that this is the approved version of the following dissertation:

Engineering and Evolution of *Saccharomyces cerevisiae* for Muconic Acid Production

Committee:

Hal Alper, Supervisor

Jeffrey Barrick

James (Jim) Bull

Marvin Whiteley

Claus Wilke

Engineering and Evolution of *Saccharomyces cerevisiae* for Muconic Acid Production

by

John Michael Leavitt, B.S. Bioch.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin
August 2016

Dedication

For my loving wife LeeAnn
and the friends and family who have supported me all these years

Acknowledgements

I would like to first thank my advisor, Hal Alper, for giving me the opportunity to engineer microorganisms and learn from someone so competent and driven. His ability to communicate science and manage a research group has set a high standard which I hope to live up to. I would next like to thank my committee members: Jeff Barrick, Jim Bull, Claus Wilke and Marvin Whiteley for their availability, their advice and especially for the practical reminder that everything will take longer than I think it will.

The Alper Lab has been a home for me and I appreciate its members for being thoughtful colleagues, helpful collaborators and wonderful friends. Kate Curran was instrumental in introducing me to molecular cloning, strain engineering and the process of scientific publication. Amanda Lanza, Johnny Blazeck and Eric Young created a supportive lab culture that has continued on since their graduation. Nathan Crook, Jie Sun, Joseph Cheng, Aaron Lin and Sun-mi Lee have been the best friends and we have shared many delightful dinners together. I owe Leqian Liu credit for helping me decide to begin the work presented in Chapter 4 and tempering my excessive optimism. Joe Abatemarco was a stalwart compatriot in sensor construction and gave an outside perspective at critical moments in my work. Kelly Markham, Haibo Li and Andrew Hill each provided significant help in setting up the bioreactor fermentation presented in Chapter 4. Heidi Redden introduced me to the art of culturing and assaying in 96-well plates, which was very helpful in finishing the work presented in Chapter 3. Matt Deaner provided a healthy challenge to my title of 2015 Alper Lab Wing-Eating Champion. Nicholas Morse and Madeline Flexer Harrison provided help refining my final oral presentation of this work. Lauren Cordova and Claire Palmer have been delightful

coworkers and I am excited to see their future accomplishments. Finally, I look forward to collaborating with James Wagner and leaving the production of muconic acid in his capable hands.

During my graduate career, I have had the opportunity to work with many fantastic undergraduate researchers and I am grateful for their help in conducting my experiments and teaching me how to lead a team. Alice Tong helped me throughout her undergraduate career, providing significant assistance in building and screening the hybrid promoters in Chapter 3 and preparing the muconic acid producing strains in Chapter 4. Joyce Tong and John Pattie contributed the work in Chapter 3 by constructing and screening strains. Cuong Ti Chi did a significant amount of the subculturing and selection work described in Chapter 4. Annie Liu helped with strain construction and screening work in Chapter 4. Sarah Ma, Inem Utin, Aditya Desai, Diem Ho and Lisa D'costa helped me figure out what could and couldn't be improved with an ARO9 based biosensor.

A major source of funding throughout my graduate career was working as a Teaching Assistant for the School of Biological Sciences. As a result, I had the opportunity to work these exceptional Professors and Instructors from whom I gained an appreciation for the art of education: Richard Meyer, Stacie Brown, Stephen Trent, Pratibha Saxena, Anita Latham, Jen Moon and Randy Linder.

And finally thank you to my family, who made this possible.

Engineering and Evolution of *Saccharomyces cerevisiae* for Muconic Acid Production

John Michael Leavitt, Ph.D

The University of Texas at Austin, 2016

Supervisor: Hal Alper

The advent of metabolic engineering and synthetic biology has resulted in a proliferation of microbial cell factories capable of producing valuable chemical products in diverse microbial hosts. This promises to provide a means to produce many of the chemical products which are currently derived from petroleum in an alternative, environmentally friendly, renewable process. Muconic acid is a chemical of particular interest for bioproduction as it can serve as a precursor for many compounds including the polymers nylon and polyethylene terephthalate. My initial research resulted in importing the biosynthetic capacity for muconic acid into the yeast host *Saccharomyces cerevisiae*. Through this work, we demonstrated the novel production of muconic acid for the first time in yeast and performed subsequent strain engineering to increase titers to 140mg/L, then the highest titer of any product from the shikimate pathway in yeast [1].

To further improve muconic acid titers, we chose to use adaptive laboratory evolution to complement initial, rational metabolic engineering efforts. To facilitate the screening of mutant strains with increased muconic acid production, a transcription-factor based biosensor was created. This biosensor was created to detect aromatic amino acids as a surrogate for flux through the shikimate pathway, the precursor pathway also used for muconic acid biosynthesis. This biosensor was based on the Aro80p transcription

factor and demonstrated both tunable induction upon aromatic amino acids as well as a constitutive mode that created ultra-strong promoters capable of two-fold stronger expression that TDH3 (GPD), one of the strongest promoters available in yeast [2].

Finally, the utility of this biosensor coupled with adaptive laboratory evolution was demonstrated in a further approach to increase muconic acid production. Namely, this sensor was used in a biosensor-enabled adaptive laboratory evolution scheme to increase titers in our original strain to over 550 mg/L muconic acid in shake flask and 1.94g/L in a fed-batch bioreactor. This work represents a 14-fold improvement in titer over our previously engineered strain and nearly a 400-fold increase over simple heterologous expression of the pathway. These results demonstrate the power of coupling rationale engineering with adaptive engineering to increase product titers.

Table of Contents

List of Tables

List of Figures

CHAPTERS

Chapter 1: Introduction and Background [1]

1.1 METABOLIC ENGINEERING OF MICROBIAL HOSTS FOR CHEMICAL PRODUCTION

> If the 20th century was the century of physics, the 21st century will be the century of biology. While combustion, electricity and nuclear power defined scientific advance in the last century, the new biology of genome research—which will provide the complete genetic blueprint of a species, including the human species—will define the next.
>
> - Craig Venter and Daniel Cohen [3]

Biology seems poised to create disruptive technologies in the chemical industry. Specifically, a wide variety of commodity chemicals have been produced with microbial hosts functioning as biocatalysts [4-8]. Moreover, these organisms facilitate the production of these molecules from renewable sources. These processes have a number of advantages over existing production streams: utilizing diverse and sustainable feedstocks, reduced emission of greenhouse gases, reduced toxic or hazardous starting materials and/or intermediates, and "green" technology designation.

While significant efforts have expanded the chemical palate which can be produced by these microorganisms [4, 7, 9, 10], complementary work has also gone into expanding the range of substrates (including lignocellulosic biomass) in an effort to reduce cost and economic viability of these processes [4, 11].

Microbial hosts have produced a number of important and interesting chemicals. Specifically, many common, endogenous, metabolites have been exploited for chemical production such as ethanol, citric acid, and lipids [5-7, 12]. Secondary metabolites, such

[1] Leavitt, J. M., Alper, H. S., Advances and current limitations in transcript-level control of gene expression. *Current opinion in biotechnology* 2015, *34*, 98-104. The author made significant contributions to preparing and editing the manuscript.

as terpenoids and other natural products have more recently been explored for microbial production[8, 13]. In more recent years, microbial hosts have been explored for the production of non-natural products that are often the result of composite pathway assembly, as is the case with muconic acid [5, 14, 15]. In these efforts, endogenous pathways can be further augmented through the integration of composite pathways from heterologous sources facilitating the formation of a diverse range of non-native products and removal of endogenous regulation [16, 17].

Heterologous expression is not enough in these host systems to create industrially-relevant titers and yields. To address these issues, a number of recently developed technologies have been explored to increase metabolic flux. Traditional metabolic engineering is the most-often attempted first step that relies on pathway modeling and the rational balancing to produce an optimal carbon flux from feedstock to product [18, 19]. Additionally, the development and maturity of systems and synthetic biology (aided by 'omics studies and gene synthesis) has expanded the number of rational changes which can be made based on model-based predictions [4, 16, 20]. Finally, advances in adaptive laboratory evolution (in ways that have progressed beyond traditional strain breeding and classical mutagenesis) have led to a further increase of pathways when the target is not known *a priori* [21-23]. Collectively, genetic manipulation such as those described above lead to improved titers, yields, and productivity leading to biocatalysts which can then be scaled-up from the bench to industrial levels [6]. These central tenants now form the basis of metabolic engineering and strain development.

1.2 BIOPRODUCTION OF MUCONIC ACID

Muconic acid, 2,3-hexadienedioic acid, is a unsaturated dicarboxylic acid that has sparked interest as a platform compound for chemical production. Muconic acid can serve as a precursor for a variety of chemical compounds including terephalic acid, adipic acid, and trimellitic acid. These three chemicals are used to synthesize polyethylene terephthalate (PET), nylon, trimellitic anhydride, other industrial plastics, food ingredients, cosmetics, and pharmaceuticals [15]. PET and nylon alone represent an addressable market value of 22 billion a year [24].

Muconic acid was initially produced in *E. coli* with continuing development as a production platform. Over 20 years, titers were raised 24-fold from 2.4 g/L to 59.2 g/L [25, 26]. While *E. coli* was initially developed as the host for industrial muconic acid production, yeast has significant advantages which make it worthy as a production host for muconic acid. Yeast has the economic advantages of lower growth temperature, lack of phage susceptibilities, less stringent nutritional requirements and the utilization of biomass byproducts as animal feeds [8, 10, 27, 28]. The low pH tolerance and ethanol production of yeast also represents an advantage for the production of an acidic compound which is more soluble in ethanol than water.

In addition to these benefits, recent work has demonstrated the use of muconic acid producing yeast in a hybrid fermentation and electrocatalytic process [24]. In that study, fermentation broth from muconic acid producing cultures was electrocatalytically hydrogenated to produce a bio-based unsaturated nylon-6,6 without requiring separation, significantly improving the economic viability of synthesizing polymer products from muconic acid producing yeast strains [24].

However, despite advances, titers remain low due to a reduced capacity to divert flux from the aromatic amino acid pathway in yeast [29-34]. Specifically, titers and yields of shikimate pathway derived molecules in yeast are lower than bacterial counterparts. Thus, additional work is necessary in the field to make yeasts a superior host for the production of these molecules. Moreover, most rational targets for this pathway have been exploited, thus requiring novel approach to further increase titers.

1.3 ADAPTIVE LABORATORY EVOLUTION

Natural evolution over long time scales resulted in all of the diversity on the planet. Researchers can harness the power of evolution in a "directed" manner to improve strains and pathways of interest. These methods, often termed "adaptive laboratory evolution" utilizes some form of screening and selection to isolate beneficial mutations [22]. Due to the global nature of these mutations and a selection scheme, this approach is most amenable to debottlenecking pathways when the targets are unexplored or unknown. This technique has been used to facilitate improved growth rates in diverse organisms and for many industrially relevant conditions such as the utilization of alternative carbon substrates and growth in the presence of toxic products and at low pH [35-41]. Selection of *S. cerevisiae* in the presence of inhibitory phenolic compounds found in lignocellulosic hydrolsate resulted in growth rate improvements of 12-57% in the presence of the inhibitors [35], while in *E. coli*, selection for growth on glycerol was able to increase conversion from glycerol to hydrogen by 20-fold [42]. However, one of the limitations associated with adaptive laboratory evolution is the ability to select for a phenotype of interest. Many desirable phenotypes, such as improvements in carbon flux

through specific metabolic paths, are actually detrimental to growth. Thus, growth selection will not work in these instances.

In situations where simple growth based selection is insufficient, researchers have turned to implementing selection strategies tailored to a chemical feature of the desired product. Examples of this are the selection of floating *Yarrowia lipolytica* cells for high lipid production [12] and resistance to hydrogen peroxide resulting in a 3-fold increase in carotenoid production in *S. cerevisiae*. The integration of complex genetic circuits which utilize riboswitches [43] or transcription factor based biosensors [21, 44] has functioned to expand the range of selectable phenotypes accessible to adaptive laboratory evolution. Through artificial selection using fluorescence activated cell sorting (FACS) in *Corynebacterium glutamicum*, Mahr and coworkers used a transcription factor based biosensor to select for a 25% improvement in L-valine titers while reducing by-product formation by 3-4 fold [44]. The ability to select for a phenotype closely related to product formation demonstrates the utility of integrating biosensors into ALE strategies.

1.4 WHOLE CELL BIOSENSORS

Whole cell biosensors provide a genetically encoded method of connecting a cell state or metabolite level to a detectable output via transcription. These can include switches utilizing Förster resonance energy transfer (FRET), RNA switches which use analyte binding to an aptamer to activate a reporter, or inducible transcription factors that change expression upon analyte concentration [45]. Collectively, biosensors offer a high throughput mechanism to screen or select for improvements in production of their respective analyte, both at the enzyme [46] and genome level [44].

As is often the case, endogenous biosensors exist based on a cellular need to regulate a given pathway. One example in *S. cerevisiae* is the gene *ARO9*, which codes for the aromatic amino acid transferase II protein [47]. This protein is the first committed step in aromatic amino acid catabolism and its expression is tightly regulated [48]. Another source of inducibility stems from the need to modulate genes expression for pathways associated with carbon consumption. Examples of catabolite-inducible gene expression include lactose [49], galactose [50] and aromatic compounds [51] . Exploring these native genetic circuits has resulted in a wide variety of tools, such as strong inducible promoters, and diverse classes of biosensors. Biosensors are not limited to their native hosts, they have been recently demonstrated that they can be developed from heterologous parts [52] as well as imported from other host systems [43]. The rapidly expanding number of biosensors represents new tools to engineer proteins, evolve strains, and build complex genetic circuits.

1.5 Control of Eukaryotic Gene Expression

1.5.1 "Parts on a Shelf" for Gene Expression

Controlling gene expression is a paramount, and often foremost, goal of most biological endeavors—from therapeutic antibody production [53] to the production of industrial enzymes [54] to the expression of heterologous metabolic pathways [4, 55]. While most of these efforts initially focus on the need for high expression, further work (especially in optimizing these processes) requires a more sophisticated, tighter control of gene expression. The need for control at many levels obviates the necessity of libraries

of synthetic parts capable of controlling transcript levels. However, not all parts are created equal and not all have been tested adequately enough to ensure function in a new system. Specifically, the current synthetic biology "parts on a shelf" model seemingly necessitates interoperability and robustness of parts, yet relies on community sourced databases to assemble experimental tools [56, 57]. This reality provides both opportunities for rapid advancement as well as a limitation in the field. These concepts are importing throughout strain engineering and biosensor development.

1.5.2 Protein Expression Control through Transcription and Translational Rates

Two major processes contribute to protein expression level: transcriptional rates and translational rates. Translation-level control (especially through tools such as ribosomal binding site calculators [58-60] and codon optimization) allow users to forward engineer the ribosomal efficiency for their gene of interest. This approach has been successfully demonstrated in prokaryotic systems where strong, orthogonal viral promoters and simpler translational mechanisms exist. In this context, translation-level control can span a 10^5 fold range [58] by editing a relatively small sequence space (such as the 5'UTR containing an RBS). Recent work on translational control in eukaryotes has focused on codon optimization to allow for improved protein expression, but the level of control of translation is not nearly as high as in prokaryotic counterparts. As an example, codon optimization of the heterologous catechol 1,2-dioxygenase gene for expression in *S. cerevisiae* resulted in a 2.9 fold increase in titer [61]. Although codon

optimization is a useful tool for yeast and higher eukaryotes, tuning transcription rates through promoters imparts a higher level of control and can achieve between a 10^2 fold dynamic range [62] and 10^4 range for orthogonal transcription factors [63]. Given the success of transcription-level control in yeast, it is important to consider both the synthetic parts that lead to control and the issue of robustness.

1.5.3 Promoters

Promoters have one of the largest impacts on gene expression and were among the first parts to be studied and diversified via random mutagenesis [64]. These initial efforts were marked by a robust definition of promoter strength taking into account dilutions by growth, the promoter's ability to impact multiple proteins, measurement of mRNA levels, and utility in heterologous pathway expression. More recent efforts aim at creating novel promoters (independent of a native scaffold) to increase the range of transcriptional capacity.

The galactose inducible promoter (GAL) is the strongest yeast inducible promoter; however it suffers from complete repression by glucose. Liang and coworkers developed a novel gene switch that coupled the inductive strength of the GAL promoter with the tight binding affinity of estradiol for the estrogen receptor protein. This ultimately led to a series of parts capable of inducing a multistep pathway using 10nM estradiol in the presence of glucose and resulting in a 50 fold improvement in zeaxanthin production over previous efforts using constitutive promoters [65].

Some of the strongest yeast promoters have been constructed through a hybrid approach by coupling upstream activating sequences (UAS) with a core promoter. Adjusting the composition of the UAS elements enables upwards of 50 to 300 fold dynamic range in expression strength, reaching some of the highest reported strength of a promoter in *S. cerevisiae* [62, 66]. Improved core promoters could lead to even greater transcriptional control in these systems. Core promoters were investigated in the yeast *Pichia pastoris* and synthetic core promoters were designed using common sequence motifs and transcription factor binding sites. These synthetic core promoters were combined with the methanol inducible promoter pAOX1 to generate diverse activity between 10% to 117% of the wild-type promoter, however only fluorescent protein expression was reported [67]. These hybrid promoter approaches represent an opportunity to "dial-in" a specific quantity of an activating sequence, producing promoters with a specific strength.

1.5.4 Trans-Acting Factors

Each of the DNA constructs described above were characterized independent of trans-acting factors that may be used to further augment transcription control. Moreover, trans-factors can be engineered to be orthogonal to the native transcriptional machinery allowing for a synthetic separation of pathways and regulation. As examples, T7 RNA polymerase variants were generated for *E. coli* that recognize unique promoter sequence

8 to 75 fold more than off target promoters leading to the ability to control multiple pathways [68]. CrisprTF's developed by Farzadhad and co-workers based on the CRISPR/Cas system from *Streptococcus pyogenes* use an endonuclease deficient Cas9 combined with an activation domain to enable up to 70-fold activation of desired promoters in HEK293T cells and *S. cerevisiae* [69, 70]. Trans-acting factors play a major part in facilitating gene expression and their engineering represents a powerful tool for creating high strength genetic circuits.

1.6 SUMMARY

Metabolic engineering and synthetic biology tools have the ability to redesign microbial genomes to establish new organisms capable of producing a diverse array of chemicals of interest. In parallel to this, adaptive laboratory evolution provides a mechanism to augment the production of these chemicals from laboratory scale to industrially relevant titers. Industrial production can be further improved by engineering strains to utilize inexpensive carbon sources. Across these efforts, biosensors can facilitate the direct measurement of metabolites of interest and represent a potential to direct adaptive laboratory evolution schemes to select for phenotypes regardless of growth rate and overall fitness of the strain. Finally, the development of tools to control eukaryotic gene expression represents an important area of research which will benefit the fields of metabolic engineering and bioproduction. Moreover, this control is required to enable both strain engineering and synthetic biology.

This dissertation uniquely couples the methodologies of rational and adaptive strain engineering for the production of muconic acid in yeast. The following chapters describe a complete story of rational engineering of a microorganism for production of chemical product, tool development, and the utilization of that tool to improve gene expression and direct the evolution of a strain capable of producing industrially relevant titers of our muconic acid. Specifically, Chapter 2 describes the first heterologous production of muconic acid in yeast utilizing a three-step composite pathway. We then demonstrate further genetic modifications using metabolic modeling and feedback inhibition mitigation to improve titers 24-fold. Chapter 3 describes the coordinated engineering of cis-acting elements in concert with a mutant trans-acting factor to develop a strong and modular expression system. This results in an expression system capable of transcriptional output two-fold higher than TDH3 (GPD), one of the strongest promoters to-date. Finally, Chapter 5 describes the utilization of the AAA inducible promoter from Chapter 4 as a biosensor in an ALE selection scheme to evolve strains of yeast capable of increased AAA production. We then re-route flux into the composite pathway and produce the four-fold higher than the previous highest titer of muconic acid production.

Chapter 2: Metabolic Engineering of Muconic Acid Production in *Saccharomyces cerevisiae*[2]

2.1 CHAPTER SUMMARY

The dicarboxylic acid muconic acid has garnered significant interest due to its potential use as a platform chemical for the production of several valuable consumer bio-plastics including nylon-6,6 and polyurethane (via an adipic acid intermediate) and polyethylene terephthalate (PET) (via a terephthalic acid intermediate). Many process advantages (including lower pH levels) support the production of this molecule in yeast. In this chapter, we present the first heterologous production of muconic acid in the yeast *Saccharomyces cerevisiae*. A three-step synthetic, composite pathway comprised of the enzymes dehydroshikimate dehydratase from *Podospora anserina*, protocatechuic acid decarboxylase from *Enterobacter cloacae*, and catechol 1,2-dioxygenase from *Candida albicans* was imported into yeast. Further genetic modifications guided by metabolic modeling and feedback inhibition mitigation were introduced to increase precursor availability. Specifically, the knockout of *ARO3* and overexpression of a feedback-resistant mutant of *aro4* reduced feedback inhibition in the shikimate pathway, and the *zwf1* deletion and over-expression of *TKL1* increased flux of necessary precursors into the pathway. Further balancing of the heterologous enzyme levels led to a final titer of nearly 141 mg/L muconic acid in a shake-flask culture, a value nearly 24 fold higher than

[2] Curran, K. A., Leavitt, J. M., Karim, A. S., Alper, H. S., Metabolic engineering of muconic acid production in *Saccharomyces cerevisiae*. *Metabolic engineering* 2013, *15*, 55-66. The author made significant contributions to designing, conducting and analyzing the experiments as well as preparing and editing the manuscript.

the initial strain. Moreover, this strain has the highest titer and second highest yield of any reported shikimate and aromatic amino acid-based molecule in yeast in a simple batch condition. This chapter collectively demonstrates that yeast has the potential to be a platform for the bioproduction of muconic acid and suggests an area that is ripe for future metabolic engineering efforts.

2.2 INTRODUCTION

Worldwide pressures to reduce petroleum footprints have increased interest in alternative, renewable methods to produce nearly all commodity and specialty chemicals. To this end, the field of metabolic engineering has begun to answer this demand through the development of organisms that can produce an increasingly diverse array of chemicals [4, 71-74]. In particular, bio-plastics have become an especially potent area as demonstrated by the metabolic engineering of strains for production of precursors such as succinic acid, ethylene glycol (from bio-ethanol), 1,3-propanediol, 1,4-butanediol, ρ-hydroxystyrene, styrene, as well as the development of novel bio-plastics such as polylactides and polyhydroxyalkanoates [75-83]. Beyond this list, muconic acid serves as another interesting precursor and platform chemical for producing several bio-plastics. Muconic acid is easily converted via hydrogenation into adipic acid, a chemical used to produce nylon-6,6 and polyurethanes. Additionally, muconic acid can be converted via the Diels-Alder reaction with acetylene and subsequent oxidation into terephthalic acid, one of two primary constituents in the plastic polyethylene terephthalate (PET). Terephthalic acid is also used in the production of polyester. World production of adipic acid and terephthalic acid is over 2.8 and 71 million tonnes, respectively [84, 85]. At present, both of these chemicals are primarily produced from non-renewable petroleum

feedstock and toxic intermediates, thus warranting a sustainable, biosynthetic production platform.

Muconic acid is not endogenously produced from carbohydrates by any known organism. However, muconic acid can be found during the catabolism and detoxification of aromatic compounds by some organisms, including yeast such as *Candida sp*., and bacteria such as *Acinetobacter sp.*, *Rhodococcus sp*., and *Sphingobacterium sp*., among others [86-89]. Previously, Draths and Frost engineered a recombinant *Escherichia coli* to produce muconic acid from glucose via a heterologous synthetic pathway drawing from a naturally occurring intermediate in the shikimate pathway, 3-dehydroshikimate (DHS) [26, 90]. In this synthetic, composite pathway, DHS is converted to protocatechuic acid (PCA) via a DHS dehydratase cloned from *Klebsiella pnemoniae*, PCA is then converted to catechol via a PCA decarboxylase from *K. pnemoniae*, and finally catechol is converted to cis,cis-muconic acid via a catechol 1,2-dioxygenase from *Acinetobacter baylyi* (**Figure 2.1**). This pathway along with some minor modifications of metabolism enabled the production of muconic acid in *E. coli*.

Many industrial biotechnological processes are moving toward using yeasts as platform organisms due to their many advantages. The yeast *Saccharomyces cerevisiae* is an ideal host organism for industrial chemical production because it offers advantages including withstanding lower temperatures, easier separations, no phage contaminations, suitability in large-scale fermentation, lower pH fermentations, and generally higher tolerances. *S. cerevisiae* has been explored as a host for producing heterologous models

that utilize precursors in the shikimate and aromatic amino acid pathways such as vanillin, ρ-hydroxybenzonic acid, ρ-amino benzoic acid, ρ-hydroxycinnamic acid, resveratrol and naringenin [29, 30, 32, 33, 91-93]. These examples and advantages raise the possibility of using yeast as a platform for the production of muconic acid. Additionally, *S. cerevisiae* naturally prefers a lower pH environment than *E. coli*, a condition better suited for producing a di-acid. Here, we present the first reported production of muconic acid in the yeast *S. cerevisiae*. Through a series of strain modifications, yields were increased more than 20-fold from the initial parental strain and resulted in the highest titer of an aromatic-based molecule in a yeast shake-flask (over 140 mg/L) and among the highest yields.

2.3 RESULTS

2.3.1 Enzyme characterization and pathway assembly

Since muconic acid is not an endogenous metabolite, it is necessary to recreate a synthetic production pathway in yeast. To create the initial pathway, we sought to utilize the same three enzymes classes employed to produce muconic acid in *E. coli* [26]. This pathway converts DHS, an intermediate in the shikimate pathway (and ultimately the aromatic amino acid biosynthesis pathways) into muconic acid in three steps (**Figure 2.1**). These three steps are carried out by a DHS dehydratase, a PCA decarboxylase, and a 1,2-catechol dioxygenase.

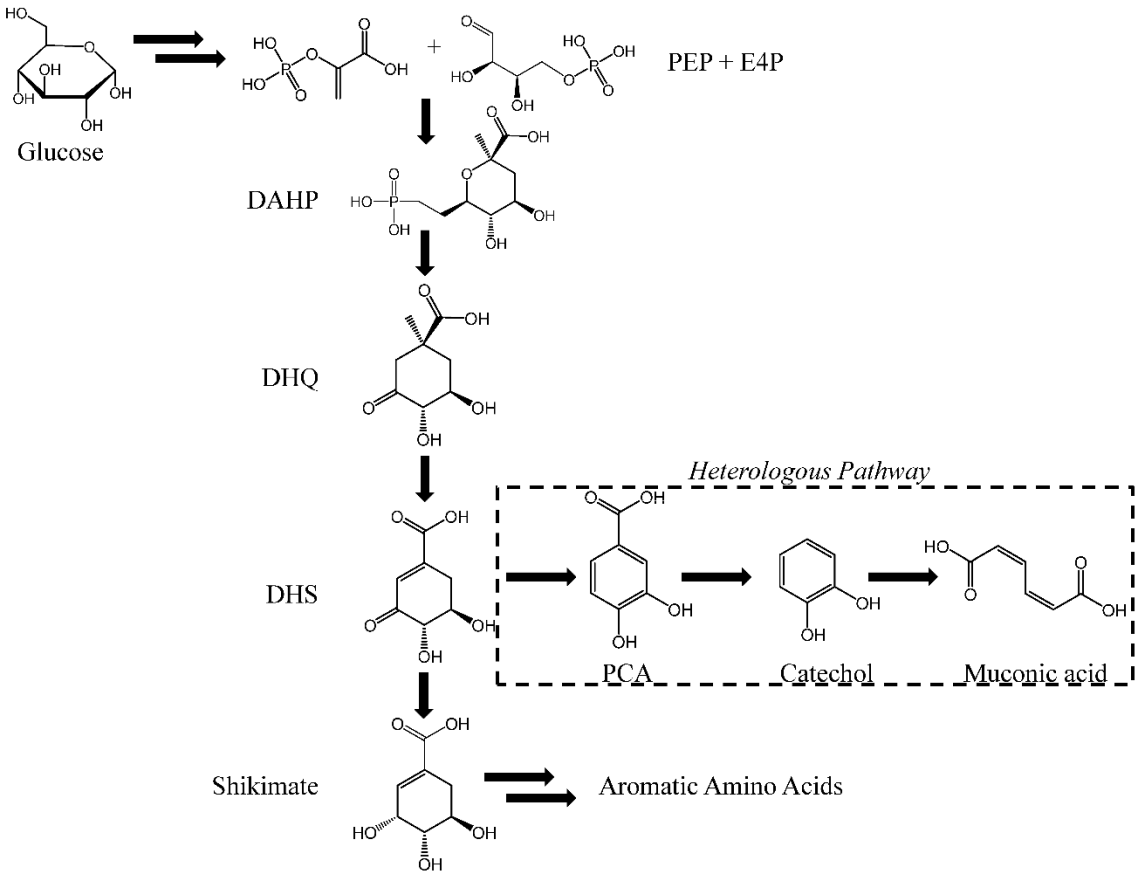

Figure 2.1 Composite heterologous pathway for muconic acid production.

The synthetic pathway for muconic acid is depicted in the context of the shikimate pathway in yeast. The following metabolite abbreviations are used: PEP is phosphoenolpyruvate, E4P is erythrose-4-phosphate, DAHP is 3-deoxy-D-arabinoheptulosonate-7-phosphate, DHQ is dehydroquinate, DHS is dehydroshikimate, and PCA is protocatechuic acid.

A two-step approach was employed to identify heterologous enzymes for this synthetic pathway. First, candidate DHS dehydratase and catechol 1,2-dioxygenase enzymes were individually expressed in *S. cerevisiae* and tested for activity using an *in*

vitro enzyme assay. These tests were straight-forward due to the availability of a spectrophotometric enzyme assay. Second, the best performing DHS dehydratase and catechol 1,2-dioxygenase were co-expressed along with each candidate PCA decarboxylase enzyme and tested for the ability to complete the pathway and produce muconic acid. This complementation type assay was conducted due to the lack of a suitable *in vitro* PCA decarboxylase enzyme activity assay.

Several DHS dehydratase enzymes have been previously characterized and studied in literature. The *AroZ* gene from *K. pneumoniae* encodes a DHS dehydratase that has previously been heterologously expressed in *E. coli* [26, 90]. However, this gene has never been expressed in *S. cerevisiae* and its function was uncertain given its bacterial origin. As a result, this gene was codon- and expression-optimized for *S. cerevisiae* and synthesized by Blue Heron Biotech to create the plasmid p413-TEF-kp*AroZ*$_{opt}$. The non-optimized version of the gene was also cloned and tested in this chapter (the resulting plasmid was named p413-TEF-kp*AroZ*). Additional *AroZ* homologues have either been identified in literature or could be selected on the basis of sequence homology. The gene Pa_5_5120 from *P. anserina* (also known as *P. pauciseta*) is a DHS dehydratase that has been successfully expressed in *S. cerevisiae* as a first step in the pathway to produce vanillin [29]. As a result, we also codon- and expression-optimized this gene and produced the plasmid p413-TEF-pa5_5120$_{opt}$. Another well-studied DHS dehydratase gene is the QutC gene from *Aspergillus sp.* [94]. This gene from *A. niger* was likewise codon- and expression-optimized and included in this chapter

as plasmid p413-TEF-an*QutC*$_{opt}$. Finally, a potential homologue from *D. hansenii* was identified via a BLAST search of the fungi kingdom using the *K. pneumoniae AroZ* gene as a search query. This gene was cloned directly from *D. hansenii* gDNA and included in this chapter as plasmid p413-TEF-dhDEHA2F15906g. These four genes (five combinations as *K. pneumoniae AroZ* was included in both codon optimized and non-optimized form) were each individually expressed in *S. cerevisiae* on a low copy plasmid using a strong TEF promoter and assayed for activity.

Transformed cells were grown and cell extracts were harvested and tested for *in vitro* DHS dehydratase activity via a spectrophotometric assay. Only two of the five DHS dehydratase constructs yielded active enzymes with detectible *in vitro* kinetic activity; the codon-optimized forms from *K. pnemoniae* and from *P. anserina*. Experimentally measured kinetic constants (K_m and V_{max}) were nearly two-fold more favorable for the *P. anserina* DHS dehydratase over the *K. pneumonia AroZ* (**Table 2.4**, at the end of the chapter). These results demonstrated the superiority of the fungal source for expression in yeast of this particular enzyme.

Next, we sought to identify a suitable catechol 1,2-dioxygenase for the muconic acid synthetic pathway. Similarly to the DHS dehydratase genes, several genes from both bacterial and fungal sources were tested on the basis of *in vitro* activity of cell lysate. First, both the wild-type and codon- and expression-optimized versions of the *CatA* gene from *A. baylyi* were tested (plasmids p413-TEF-ab*CatA* and p413-TEF-ab*CatA*$_{opt}$). The wild-type version of this gene was previously used in *E. coli* for the

assembly of a muconic acid pathway [26, 90]. Next, the *HQD2* gene from *C. albicans* [95] was codon- and expression-optimized and expressed in plasmid p413-TEF-ca*HQD2*$_{opt}$. Finally, a homologue to the *CatA* gene was found in *D. hansenii* using a BLAST search. This gene was cloned directly from gDNA and expressed in plasmid p413-TEF-dhDEHA2C14806g.

Unlike the *AroZ* selection, all of the four putative catechol 1,2-dioxygenase were proven to be active on the basis of *in vitro* activity assays (**Table 2.5**, at the end of the chapter). However, the enzyme kinetics did not immediately point toward the superiority of one particular *CatA* gene and thus an additional *in vivo* feeding assay was conducted. To do so, 1 g/L catechol was added to stationary phase cultures expressing each catechol 1,2-dioxygenase enzyme and supernatants were assayed for muconic acid using HPLC after 24 hours of culture. Using this feeding assay, the *HQD2* gene from *C. albicans* produced the largest amount of muconic acid (**Figure 2.2**) and was selected as the candidate gene moving forward. Moreover, in this assay, we were able to account for all 1 g/L of catechol after 24 hours as either free catechol or muconic acid product, thus precluding any major degradation products formed in this timeframe.

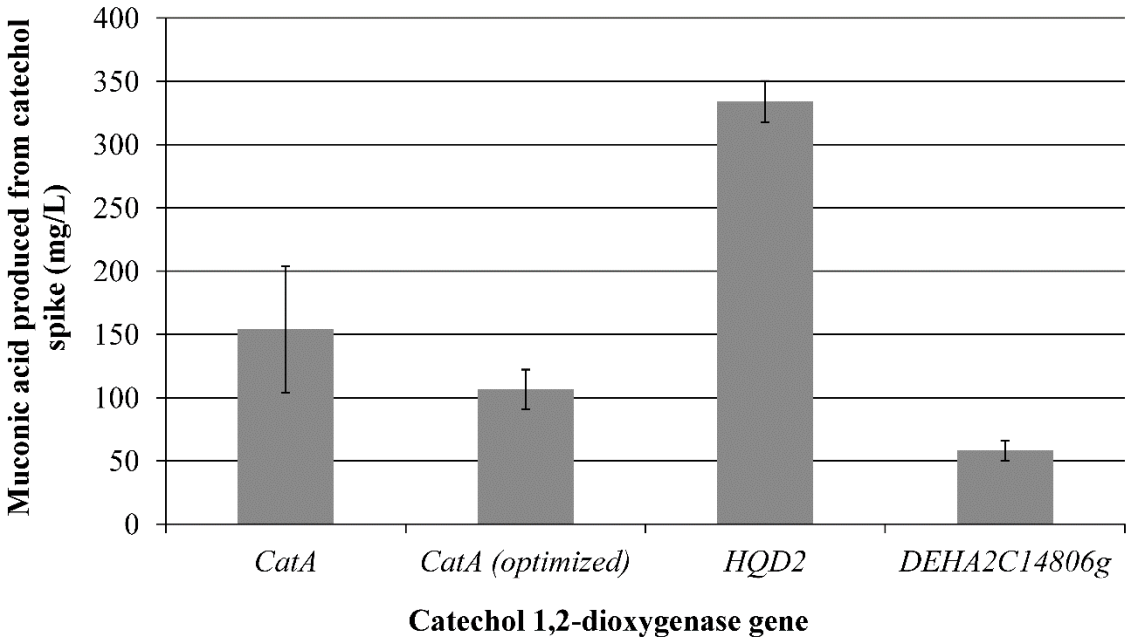

Figure 2.2 Production of muconic acid from catechol 1,2-dioxygenase-expressing strains.

To test for *in vivo* function of the catechol 1,2-dioxygenases, muconic acid concentration in the culture supernatant was measured using HPLC 24 hours after flasks were spiked with 1 g/L catechol. Standard deviation values are based on results from biological triplicates.

Finally, once the DHS dehydratase and catechol 1,2-dioxygenase enzymes were chosen for the yeast synthetic muconic acid pathway, several PCA decarboxylase candidates were characterized. The *AroY* gene from *K. pneumoniae* was codon- and expression-optimized and was expressed in plasmid p416-TEF-kp$AroY_{opt}$. The wild-type gene, which has previously been expressed in *E. coli* [26, 90], was also characterized using plasmid p416-TEF-kp*AroY*. The gene ECL_01944 from *E. cloacae* was also codon- and expression-optimized and expressed in plasmid p416-TEF-ECL_01944_{opt}

[96, 97]. Additionally, a BLAST search was used to look for possible *K. pneumoniae AroY* homologues in the fungal kingdom. The DEHA2G00682g gene from *D. hansenii* was identified and cloned from gDNA (plasmid p416-TEF-dhDEHA2G00682g). Additionally, we hypothesized that the *FDC1* and *PAD1* genes from *S. cerevisiae*, which together are known to be phenylacrylate decarboxylases [98], may have some PCA decarboxylase activity. As a result, these two genes were cloned and co-expressed on a single plasmid (plasmid p416-TEF-scFDC1/PAD1). Finally, the annotated genome for *P. anserina* lists two 2,3-dihydroxybenzoic acid decarboxylases, Pa_0_880, and Pa_4_4540 (PCA is 3,4-dihydroxybenzoic acid). We hypothesized that one of these could also have promiscuous PCA decarboxylase activity and therefore codon- and expression-optimized these genes and cloned them in plasmids p416-TEF-pa0_880$_{opt}$ and p416-TEF-pa4_4540$_{opt}$. These six different genes (seven total variations) were transformed into yeast and tested using an *in vivo* pathway completion assay due to the lack of a simple *in vitro* enzymatic assay for the PCA decarboxylase.

Each of the PCA decarboxylase candidates was expressed in a strain that also harbored plasmids containing the DHS dehydratase gene from *P. anserina* and the *HQD2* gene from *C. albicans* (p413-TEF-pa5_5120$_{opt}$ and p415-GPD –ca*HQD2*$_{opt}$) and tested for muconic acid production using HPLC. Strains were cultivated in shake flask conditions with a starting OD_{600} of 0.25. Muconic acid concentration was quantified in the culture supernatant after 48 hours of culture. Only the PCA decarboxylases from *K. pnemoniae* and *E. cloacae* were active, and the two codon-optimized enzymes had

similar production values (**Table 2.6**, at the end of the chapter). The strain with *AroY* from *K. pnemoniae* was named MuA01 and was used first in this chapter. It was discovered that the *E. cloacae* gene performed better in more optimized strains and was used strains developed later in this chapter. Not only did this experiment identify active PCA decarboxylase enzymes, it also represented the first time that muconic acid has been successfully produced in *S. cerevisiae*. Titers were very low (around 5 mg/L) indicating that more extensive strain and pathway engineering was necessary.

2.3.2 Relief of Amino Acid Feedback Repression

Aromatic amino acid biosynthesis is a tightly regulated process with the flux of intermediates in the shikimate and aromatic pathways subject to significant allosteric regulation [99]. To determine the magnitude of impact in our strain, we assayed the effect of exogenous repression by removing amino acid supplementation in the media. Culturing strain MuA01 in a synthetic minimal media (YSM) lacking the exogenous aromatic amino acids was shown to increase muconic acid production three-fold over a synthetic complete media (YSC) (**Figure 2.3**). These results demonstrate that feedback inhibition caused by the aromatic amino acids in the media was limiting flux to the muconic acid pathway and also implicated that intracellular endogenous production may be further limiting this pathway.

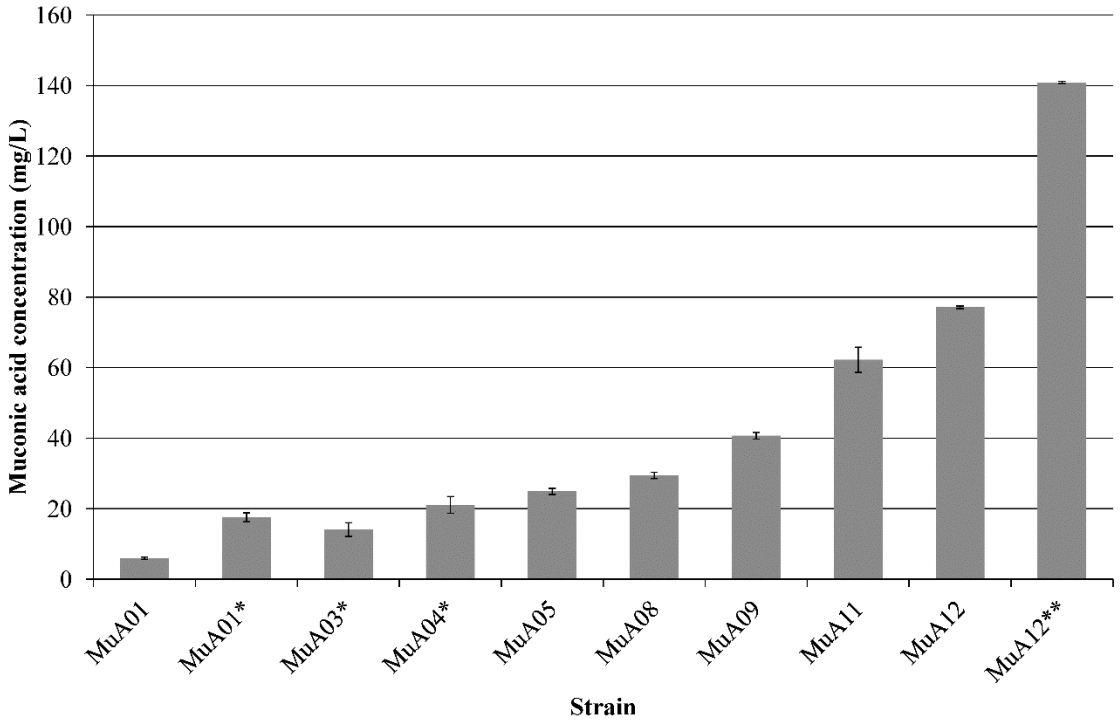

Figure 2.3 Muconic acid production across strains used in this chapter.

Muconic acid levels are provided in the progression from first pathway assembly to the final, optimized strain. These strain names are described in more detail with genotypes in **Table 2.2**. All values were determined by HPLC after at least 48 hours of batch flask culture. Strains were grown in complete synthetic media (YSC) unless marked as follows: *Strains grown in minimal synthetic media (YSM), **Strains grown in complete synthetic media (YSC) with 40 g/L glucose. Standard deviations are based on biological triplicates.

We next removed known feedback inhibition through genetic modification. Entry into the shikimate pathway from the pentose phosphate and glycolytic pathways is governed by 3-deoxy-D-arabinoheptulosonate7-phosphate (DAHP) synthase, an enzyme that catalyzes the condensation of phosphoenolpyruvate and erythrose-4-phosphate to DAHP. Yeast has two isozymes of DAHP synthase that are regulated independently by

phenylalanine (*ARO3*) and tyrosine (*ARO4*) feedback inhibition. It has been previously shown that knocking out both *ARO3* and *ARO4* and over-expressing a mutant version, $aro4_{k229l}$, can alleviate the feedback inhibition in this step [100]. Therefore, serial gene deletion was performed to obtain an *aro3 aro4* double knockout the BY4741 strain, resulting in strain MuA02. Next, a plasmid containing the mutant $aro4_{k229l}$ (p416-TEF-$aro4_{k229l}$) was transformed along with the three muconic acid pathway genes (on plasmids p415-GPD-ca*HQD2*opt and p413-TEF-pa5_5120_{opt}/TEF-kp*AroY*$_{opt}$). For comparison, the wild type *ARO4* gene was also over-expressed on plasmid p416-TEF-sc*ARO4* in the same background. The strain with *ARO4* was named MuA03 and the strain with $aro4_{k229l}$ was named MuA04. The strain expressing $aro4_{k229l}$ achieved a 50% increase in muconic acid production over wild-type *ARO4*, producing 21 ± 2 mg/L and 14 ± 2 mg/L, respectively (**Figure 2.3**). Furthermore, unlike the wild-type *ARO4*, the strain expressing $aro4_{k229l}$ produced the same amount of muconic acid in both YSC and YSM, demonstrating that the feedback inhibition in the pathway had been removed. Subsequently, $aro4_{k229l}$ was integrated into the *ARO4* genomic locus under control of the GPD promoter (strain MuA06).

2.3.3 Over-expression and balancing of heterologous pathway enzymes

Due to the fact that the muconic acid pathway is a synthetic, composite pathway comprised of enzymes from several different sources, the enzyme activities and expression levels require proper balancing. When pathway intermediate concentrations

were measured in addition to muconic acid, it was immediately apparent that PCA decarboxylase is an important rate limiting step in the pathway. For example, in strain MuA04, the PCA concentration reached more than seven times the level of muconic acid, to 166 ± 11 mg/L. To address this bottleneck, the PCA decarboxylase gene was changed from the *K. pneumoniae AroY*$_{opt}$ to ECL_01944$_{opt}$ from *E. cloacae*, an enzyme with a slightly higher initial muconic acid production in enzyme evaluations described above (**Table 2.6**, at the end of the chapter). ECL_01944$_{opt}$ was also cloned into a high copy 2μ plasmid (plasmid p425-GPD-ECL_01944$_{opt}$) instead of the centromeric plasmid originally used (strainMuA05). This change resulted in an increase of muconic acid production to 25 ± 1 mg/L (**Figure 2.3**). The PCA decarboxylase gene was subsequently further over-expressed by integrating it into the Ty2 retrotransposon δ elements multiple times under the control of the GPD promoter using the pITy3 vector [101] (strain MuA07). This integration did not appreciatively increase the production of muconic acid although it did increase the mRNA expression of the gene roughly 3-fold (**Figure 2.4**). As a result, it is likely that the problems associated with PCA decarboxylase are post-transcriptional in nature. In the final muconic acid producing strain, the PCA decarboxylase was over-expressed using both multiple genomic integrations and a high copy plasmid (Subchapter 2.3.5). Finally, the other two pathway enzymes, DHS dehydratase and catechol 1,2-dioxygenase, were over-expressed on high-copy plasmids to further improve the pull of metabolites through the synthetic pathway (strain MuA08). Collectively, these changes increased the production of muconic acid to 30 ± 1 mg/L muconic acid (**Figure 2.3**).

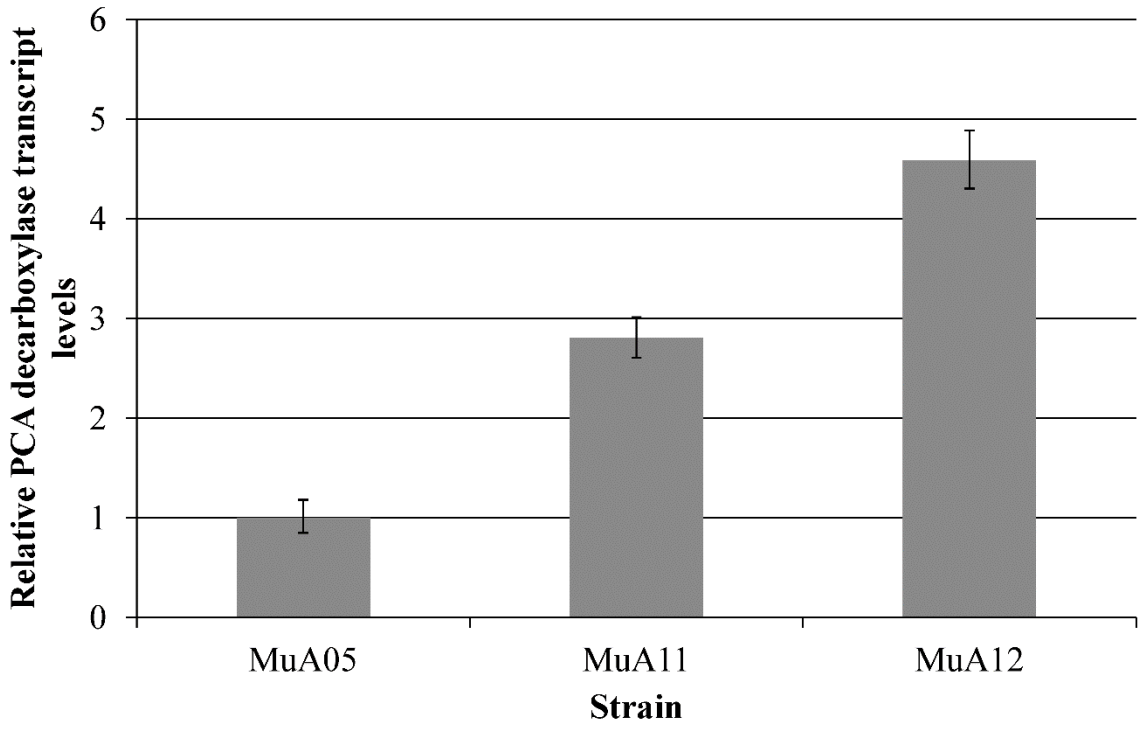

Figure 2.4 Transcript levels for PCA decarboxylase overexpression strains.

Select strains with overexpression of the PCA decarboxylase gene ECL_01944$_{opt}$ either by multi-copy plasmid or multi-copy integration were measured by RT-PCR. Strains are described in **Table 2.2**. Transcript levels increased with successive genetic engineering. Standard deviation values are based on three technical replicates.

2.3.4 Flux balance analysis allows for further improvements in precursor availability

To identify additional targets for metabolic engineering of muconic acid in *S. cerevisiae*, we next utilized the framework of flux balance analysis. A similar approach has been previously used to improve metabolite production for a variety of products [102-108]. The genomic model iMM904 [109] was used as a starting point and additional reactions

were included to account for the heterologous muconic acid pathway (**Table 2.3**, at the end of the chapter). In order to calculate the maximum theoretical yield, the system of linear equations was solved while maximizing the reaction for muconic acid production. This resulted in a value of 85.7% mol/mol from glucose. It was immediately noted, however, that this solution required maximizing the flux of fructose-6-phosphate and glyceraldehyde-3-phosphate through the transketolase reaction in the pentose phosphate pathway to produce erythrose-4-phosphate and xylulose-5-phosphate. This flux mode balances the availability of erythrose-4-phosphate and phosphoenolpyruvate and avoids the oxidative shunt of the pentose phosphate pathway. However, it is unlikely that this flux mode occurs endogenously *in vivo* due to kinetic constraints [110]. It is more likely that flux enters into the pentose phosphate pathway through the glucose-6-phosphate dehydrogenase reaction and that the transketolase reaction utilizes erythrose-4-phosphate and xylulose-5-phosphate to produce fructose-6-phosphate and glyceraldehydes-3-phosphate (the reverse of what is desired). When we forced flux to go toward this route in the *in silico* model, the maximum theoretical yield was decreased to 60.9% mol/mol glucose. These results suggest the need for a genetic modification to rewire the pentose phosphate pathway flux.

In order to implement this desired flux network *in vivo*, two genetic steps were taken. First, the transketolase gene, *TKL1*, was over-expressed on p413-TEF-sc*TKL1* (strain MuA09) to help favor the kinetically hindered pathway. Second, the glucose-6-phosphate dehydrogenase gene, *ZWF1*, was knocked out (strain MuA10) to force entry

into the pentose phosphate pathway to occur via transketolase. When these modifications were combined in the same strain (strain MuA11), the muconic acid titer increased two-fold over the previous best strain, to a value of 62 ± 4 mg/L (**Figure 2.3**).

2.3.5 Final muconic acid-producing strain characterization

To further increase the muconic acid production in the strain developed above, the PCA decarboxylase gene was over-expressed on a high copy plasmid (p426-GPD-ECL_01944$_{opt}$) in addition to being integrated multiple times onto the chromosome (strain MuA12). This increased the PCA decarboxylase gene RNA expression by 64% over the previous strain (**Figure 2.4**), and increased the muconic acid production to 77 ± 1 mg/L (**Figure 2.3**) with a yield of 3.9 mg/g glucose. Finally, we modified the glucose content of the medium by growing MuA12 in YSC media with 40 g/L glucose supplementation for an extended period of 108 hr (**Figure 2.5**) after initial seeding at an OD_{600} of 0.25. The final muconic acid titer in this strain was 141 ± 1 mg/L. This strain produced the highest amount of muconic acid in this chapter (nearly 24 times the value produced by the initial strain) and represents the highest titer of an aromatic-based molecule produced in yeast in a simple shake-flask condition to date (**Table 2.7**, at the end of the chapter).

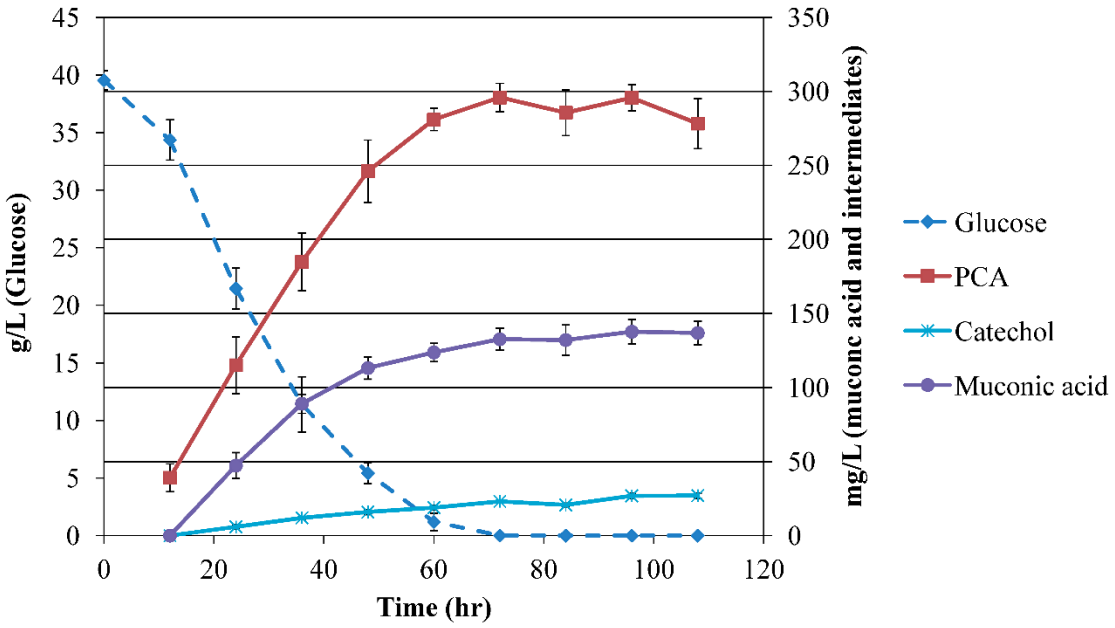

Figure 2.5 Fermentation profile of final muconic acid strain MuA12.

The concentration of muconic acid, PCA, catechol, and glucose in the culture supernatant was measured over time from a shake-flask experiment. Muconic acid levels reached the highest titer reported for an aromatic based molecule of nearly 141 mg/L. Glucose concentrations were measured using the YSI bioanalyzer, all others were measured using HPLC. Glucose values are plotted on the left axis while remaining metabolites are graphed on the right axis. Standard deviations are based on results from biological triplicates.

2.3 Discussion

This chapter reports the first successful heterologous production of muconic acid in the yeast *Saccharomyces cerevisiae*. To accomplish this production, we assembled a synthetic, composite pathway comprised of three distinct enzymes. The DHS dehydratase gene from *P. anserina* was easily identified as the best among the candidate

enzymes tested. This finding is corroborated with a previous report demonstrating successful use of this enzyme in *S. cerevisiae* for the production of vanillin [29]. In contrast, none of the catechol 1, 2-dioxygenase candidates demonstrated a difference in catalytic activity in the first *in vitro* assay. However, it became clear after an *in vivo* feeding assay that the gene from *C. albicans* had a higher capacity. It is also interesting to note that for the *K. pnuemoniae CatA* gene, the un-optimized form showed better activity than the codon-optimized form of the gene. This finding is an interesting result that challenges the need to always codon-optimize heterologous genes. Finally, the second step of the pathway, the PCA decarboxylase, proved the most difficult to identify and still remains the bottleneck of the pathway (as evinced by the high PCA concentration in **Figure 2.5**). Very few PCA decarboxylase genes have been studied in detail and only the *AroY* from *Klebsiella* species and ECL_01944 from *Enterobacter cloace* have even been sequenced [97, 111]. Furthermore, to our knowledge, no PCA decarboxylase has been identified in eukaryotes. In this chapter, we tested several potential eukaryotic enzymes, but unfortunately none possessed the desired activity. Consequently, it was not surprising to discover that PCA decarboxylase remains as a rate limiting step in this pathway. High over-expression of the PCA decarboxylase gene ECL_01944_{opt} partially relieved the pathway bottleneck, but future efforts to engineer this enzyme for better activity in *S. cerevisiae* would be beneficial.

Despite challenges in selecting high flux-supporting enzymes for this synthetic, composite pathway, we achieved the highest titer of an aromatic-based or shikimate

based molecule in yeast in a simple, batch shake-flask condition. Moreover, the maximum yield we achieved of 3.86 mg/g glucose was the second highest among all published reports (**Table 2.7**, at the end of the chapter). Any success in surpassing these values has been achieved by significant optimization of culturing conditions in a bio-reactor, such as for the production of vanillin [102], or in the conditioning and optimizing of an industrial strain, such as in the production of resveratrol [112]. These results highlight a remaining, significant challenge for yeast metabolic engineering. Indeed the greatest gains in titer achieved in this chapter were due to genetic alterations in the upstream pathway. Specifically, over-expression of the feedback resistant $aro4_{k229l}$ (strain MuA3) increased production more than three-fold over the initial strain, MuA01. Furthermore, rewiring of the pentose phosphate pathway to avoid the oxidative shunt (strain MuA11) increased production two-fold (over MuA08). Yet, total net flux through the shikimate pathway in yeast is limited. There are additional potential targets for improving this heterologous pathway. Additional gene knockout targets have been suggested by flux balance analysis, including some (such as Δ*pdc1)* that have successfully increased the production of vanillin in *S. cerevisiae* [102]. Beyond gene deletions, it is possible to also augment the shikimate pathway to divert more DHS into the synthetic muconic acid pathway. Specifically, the *ARO1* gene, a penta-functional enzyme that catalyzes the majority of the steps in the shikimate pathway [113, 114], may be altered to help reduce the flux of dehydroshikimate into shikimate, thereby shifting the balance of production from the aromatic amino acids to muconic acid. It is interesting to

note that this penta-functional enzyme is unique to yeast as compared to *E. coli* in which each of these enzymatic functions is encoded by a separate gene. As such, knockout of the *aroE* gene in *E. coli* proved to be simple method to improve muconic acid production [90].

Despite the potential targets described above, it is clear that aromatic amino acid based production in yeast is an outstanding metabolic engineering challenge. In contrast, straight-forward genetic modifications such as alleviating feedback inhibition in *E. coli* can result in high gram per liter levels of aromatic amino acids and their related products [115-117]. Indeed, muconic acid can be produced at such levels in *E. coli* as well [26, 90]. The inability to achieve higher production values from rational metabolic engineering techniques in yeast suggests that flux in the shikimate and aromatic amino acid pathways is highly regulated, likely through both global and local transcription machinery. A comprehensive 'omics analysis of amino acid production in yeast [118] demonstrates significant allosteric and transcriptional regulation throughout the various amino acid pathways, partially controlled by factors such as Gcn4p. Further analysis of the improved strains here can identify similar regulatory proteins that are responsible for controlling overall flux through the shikimate pathway. Once these targets are identified, techniques such as global Transcription Machinery Engineering [119, 120] can be used to further increase the yield and titer in *S. cerevisiae*. A recent report has demonstrated that the application of gTME using global regulatory factors can improve L-tyrosine production in *E. coli* [121].Similar improvements would ultimately aid in making yeast a

suitable platform for shikimate-based molecules especially when coupled with the process advantages and lower pH tolerance that yeast possesses over *E. coli*.

2.4 CONCLUSION

In summary, this chapter provides the first demonstration that muconic acid can be produced in *S. cerevisiae* and that metabolic engineering can be used to increase production titers by nearly 24 fold over the initial strain. This chapter presents a strain with the highest titer and second highest yield of any shikimate and aromatic amino acid-based molecule in yeast in a simple batch condition. Beyond this proof-of-concept, the genetic alterations utilized here create an excellent starting point from which additional metabolic engineering, strain development, and global regulatory engineering can be performed in order to further increase titer and yield. These results demonstrate an unanswered challenge of metabolic engineering in yeast. Nevertheless, these results demonstrate that yeast can serve as a host organism for the production of sustainable bio-plastics and polymer precursors. To facilitate the selection of further improvements in muconic acid production, our next step is to engineer a biosensor which can detect the downstream products of the shikimate pathway as a surrogate for muconic acid.

Plasmid Name	Characteristics	Primers
p413-TEF-kp*AroZ*	DHS dehydratase from *K. pnemoniae*	Fwd: GACTAGTATGGTGCGCTCTATCGCCAC Rev: CCATCGATCTAACAATACTGCATCGCCGCC
p413-TEF-kp*AroZ*$_{opt}$	Codon-optimized DHS dehydratase from *K. pnemoniae*	N/A
p413-TEF-an*QutC*$_{opt}$	Codon-optimized DHS dehydratase from *A. niger*	N/A
p413-TEF-pa5_5120$_{opt}$	Codon-optimized DHS dehydratase from *P. anserina*	Fwd: GGCGCTACTAGTATGCCATCTAAGTTAGCAATTACG Rev: GCGAATTCCTAAAGGGCAGCACTTAAT
p413-TEF-dhDEHA2F15906g	Potential DHS dehydratase from *D. hansenii*	Fwd: ACTAGTATGGCTGATTATATGGAATGC Rev: GAATTCTTATTCTTGGTTAAGTGTCAAT
p416-TEF-kp*AroY*	PCA decarboxylase from *K. pnemoniae*	Fwd: GACTAGTATGACCGCACCGATTCAG Rev: CCATCGATCGCTACCCTGGTTTTTTTCC
p416-TEF-kp*AroY*$_{opt}$	Codon-optimized PCA decarboxylase from *K. pnemoniae*	N/A
p416-TEF-dhDEHA2G00682g	Potential PCA decarboxylase from *D. hansenii*	Fwd: GGACTAGTATGAGCAATTTAAGACCAGAG Rev: CCGGAATTCCTATTTATATCCGTACGCAG
p416-TEF-pa0_880$_{opt}$	Codon-optimized potential PCA decarboxylase from *P. anserina*	Fwd: ACTAGTATGGCATTACCAGCAGAAG Rev: GTCGACTTATTTTAAACGTAGTAATGCTTT
p416-TEF-pa4_4540$_{opt}$	Codon-optimized potential PCA decarboxylase from *P. anserina*	Fwd: ACTAGTATGCTTGTTTTTGGCCA Rev: GTCGACTTATGGGTCTAATGGAAG
Table 2.1: continued next page.		

Plasmid Name	Characteristics	Primers
p416-TEF-sc*FDC1*/*PAD1*	Potential PCA decarboxylases from *S. cerevisiae*, both expressed with TEF promoter	FDC1: Fwd: GCTCTAGAATGAGGAAGCTAAATCCAG Rev: CCATCGATTTATTTATATCCGTACCTTTTCC PAD1: Fwd: GACTAGTATGCTCCTATTTCCAAGAAGAA Rev: GGAATTCTTACTTGCTTTTTATTCCTTCCC To insert the second gene into the PstI site: Fwd: AACTGCAGGGAGCTCATAGCTTCAAAATGTTTC Rev: AACTGCAGGGCCGCAAATTAAAGCCTT
p416-TEF-ECL_01944$_{opt}$	Codon-optimized PCA decarboxylase from *E. cloacae*	Fwd: GGCGCTACTAGTATGCAAAACCCAATAAATGAT Rev: ACGCGTCGACCTATTTTTTGTCAGAAAATAATTC
p413-TEF-ab*CatA*	Catechol 1,2-dioxygenase from *A. baylyi*	Fwd: GACTAGTATGGAAGTTAAAATATTCAATACTCAG Rev: CCATCGATTTACACCGCTAGACGTGG
p413-TEF-ab*CatA*$_{opt}$	Codon-optimized catechol 1,2-dioxygenase from *A. baylyi*	N/A
p413-TEF-dhDEHA2C14806g	Potential catechol 1,2-dioxygenase from *D. hansenii*	Fwd: GCTCTAGAATGGATCAAGGCTTTACAGAC Rev: CCATCGATTCAACTAGCAGCAGTAGCAG
p413-TEF-ca*HQD2*$_{opt}$	Codon-optimized catechol 1,2-dioxygenase from *C. albicans*	Fwd: GGCGCTACTAGTATGTCACAAGCTTTTACAGAATCAG Rev: CAAAGCTCGAGCTATAACTTAATTTCGGCGTCTTGT
Table 2.1: continued next page.		

Plasmid Name	Characteristics	Primers
p415-GPD-ca*HQD2*$_{opt}$	Codon-optimized catechol 1,2-dioxygenase from *C. albicans*	Fwd: GGCGCTACTAGTATGTCACAAGCTTTTACAGAATCAG Rev: CAAAGCTCGAGCTATAACTTAATTTCGGCGTCTTGT
p416-TEF-sc*ARO4*	*ARO4* gene from *S. cerevisiae*	Fwd: CTAGACTAGTATGAGTGAATCTCCAATGTTC Rev: CGTTACATATATCATTAAAAAAACATATCGATCTA
p416-TEF-sc*aro4*$_{k229l}$	Mutated *aro4* gene from *S. cerevisiae*	Primers for site directed mutagenesis: Fwd: CATTCTCACCATTTCATGGGTGTTACTCTGCATGGTGTTGCTG Rev: CAGCAACACCATGCAGAGTAACACCCATGAAATGGTGAGAATG
P416-GPD- sc*aro4*$_{k229l}$	Mutated *aro4* gene from *S. cerevisiae*	Fwd: CTAGACTAGTATGAGTGAATCTCCAATGTTC Rev: CGTTACATATATCATTAAAAAAACATATCGATCTA
P425-GPD-ECL_01944$_{opt}$	Codon-optimized PCA decarboxylase from *E. cloacae* expressed with the GPD promoter	Fwd: GGCGCTACTAGTATGCAAAACCCAATAAATGAT Rev: ACGCGTCGACCTATTTTTTGTCAGAAAATAATTC
P426-GPD-ECL_01944$_{opt}$	Codon-optimized PCA decarboxylase from *E. cloacae* expressed with the GPD promoter	Fwd: GGCGCTACTAGTATGCAAAACCCAATAAATGAT Rev: ACGCGTCGACCTATTTTTTGTCAGAAAATAATTC
Table 2.1: continued next page.		

Plasmid Name	Characteristics	Primers
p413-TEF-pa5_5120$_{opt}$/TEF-kp*AroY*$_{opt}$	Codon optimized DHS dehydratase from *P. anserina* and PCA decarboxylase from *K. pnemoniae,* both expressed with TEF promoter	Fwd: GGCGCTACTAGTATGCCATCTAAGTTAGCAATTACG Rev: GCGAATTCCTAAAGGGCAGCACTTAAT To insert the second gene into the PstI site: Fwd: AACTGCAGGGAGCTCATAGCTTCAAAATGTTTC Rev: AACTGCAGGGCCGCAAATTAAAGCCTT
p413-TEF-pa5_5120$_{opt}$/GPD-ca*HQD2*$_{opt}$	Codon optimized DHS dehydratase from *P. anserina* and Codon-optimized catechol 1,2-dioxygenase from *C. albicans*	Fwd: GGCGCTACTAGTATGCCATCTAAGTTAGCAATTACG Rev: GCGAATTCCTAAAGGGCAGCACTTAAT To insert the second gene into the PstI site: Fwd: AACTGCAGAGTTTATCATTATCAATACTCGCCATTTC Rev: AACTGCAGGGCCGCAAATTAAAGCCTT
p425- TEF-pa5_5120$_{opt}$/GPD-ca*HQD2*$_{opt}$	Codon optimized DHS dehydratase from *P. anserina* and Codon-optimized catechol 1,2-dioxygenase from *C. albicans*	Fwd: GGCGCTACTAGTATGTCACAAGCTTTTACAGAATCAG Rev: CAAAGCTCGAGCTATAACTTAATTTCGGCGTCTTGT To insert the second gene into the SacI site: Fwd: TGACTGAGCTCATAGCTTCAAAATGTTTCTACTC Rev: CAAAGAGCTCCAAATTAAAGCCTTCGAG
P413-TEF-sc*TKL1*	*TKL1* gene from *S. cerevisiae* expressed with the TEF promoter	Fwd: GCTCTAGAATGACTCAATTCACTGACATTG Rev: ACGCGTCGACTTAGAAAGCTTTTTTCAAAGGAG
pITy-GPD-ECL_01944$_{opt}$	Ty2 integration vector containing the ECL_01944$_{opt}$ gene from *E. cloacae* expressed with the GPD promoter	Fwd: TGACTGAGCTCAGTTTATCATTATCAATACTCGC Rev: GATGGTACCCAAATTAAAGCCTTCGAG

Table 2.1: continued next page.

Plasmid Name	Characteristics	Primers
pUG6	Plasmid containing *KanMX* gene with *loxP* sites	To generate *ARO3* knockout cassette: Fwd: CTACTACCCCTATTACGTTACAAGAACACTTT ATAGCATTCAGCTGAAGCTTCGTACGC Rev: TATCATTCAAGATTATTTGCATTTTTCCCTCAT TTACAGGGCATAGGCCACTAGTGGATCTG To generate *ARO4* knockout cassette: Fwd: TTTAACCGCTAAATTTAGTAAACAAAAGAATC TATCAGAACAGCTGAAGCTTCGTACGC Rev: GAGGAAAGAATGTACGTTACATATATCATTAA AAAAACATGCATAGGCCACTAGTGGATCTG To generate *ZWF1* knockout cassette: Fwd: AAGAGTAAATCCAATAGAATAGAAAACCACA TAAGGCAAGCAGCTGAAGCTTCGTACGC Rev: TTCAGTGACTTAGCCGATAAATGAATGTGCTT GCATTTTTGCATAGGCCACTAGTGGATCTG Primers for PCR confirmation of knockouts: KanMX: Fwd: CAGCTGAAGCTTCGTACGC Rev: GCATAGGCCACTAGTGGATCTG ARO3 WT: Fwd: ATGTTCATTAAAAACGATCACGC Rev: CTATTTTTTCAAGGCCTTTCTTCTG ARO4 WT: Fwd: ATGAGTGAATCTCCAATGTTCG Rev: CTATTTCTTGTTAACTTCTCTTCTTTGTCTGA ZWF1 WT: Fwd: ATGAGTGAAGGCCCCGTC Rev: CTAATTATCCTTCGTATCTTCTGGCTTAGT
Table 2.1: continued next page.		

Plasmid Name	Characteristics	Primers
pSH47	Plasmid containing Cre recombinase under control of *GAL1* promoter for removal of KanMX from integration cassettes	None

Table 2.1 Plasmids used in this chapter.

All plasmids were made from plasmids described in [122] with the exception of pITy-GPD- ECL_01944opt from the pITy3 vector [101]and pUG6 and pSH47 [123]. Primers marked N/A correspond to plasmid inserts that were obtained by restriction digest and gel extraction.

Strain	Genotype	Plasmids	Parent Strain	Reference
BY4741	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0*	none		Euroscarf Y00000
MuA01	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0*	p416-TEF- kpAroY$_{opt}$, p413-TEF-pa5_5120$_{opt}$, p415-GPD-ca*HQD2*$_{opt}$	BY4741	This Chapter
MuA02	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ*	none	BY4741	This Chapter
MuA03	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ*	p415-GPD-ca*HQD2*$_{opt}$, p413-TEF-pa5_5120$_{opt}$/TEF-kpAroY$_{opt}$, p416-TEF-sc*ARO4*	MuA02	This Chapter
MuA04	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ*	p415-GPD-ca*HQD2*$_{opt}$, p413-TEF-pa5_5120$_{opt}$/TEF-kpAroY$_{opt}$, p416-TEF-sc*aro4*$_{k229l}$	MuA02	This Chapter
MuA05	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ*	p425-GPD-ECL_01944$_{opt}$, p413-TEF-pa5_5120$_{opt}$/GPD-ca*HQD2*$_{opt}$, p416-TEF-sc*aro4*$_{k229l}$	MuA02	This Chapter
MuA06	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::P$_{GPD}$-aro4$_{k229l}$*	none	MuA02	This Chapter
MuA07	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::P$_{GPD}$-aro4$_{k229l}$; Ty2δ::P$_{GPD}$-ECL_01944$_{opt}$*	none	MuA06	This Chapter
MuA08	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::P$_{GPD}$-aro4$_{k229l}$; Ty2δ::P$_{GPD}$-ECL_01944$_{opt}$*	p425-TEF-pa5_5120$_{opt}$/GPD-ca*HQD2*$_{opt}$, p413-TEF	MuA07	This Chapter
MuA09	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::P$_{GPD}$-aro4$_{k229l}$; Ty2δ::P$_{GPD}$-ECL_01944$_{opt}$*	p425-TEF-pa5_5120$_{opt}$/GPD-ca*HQD2*$_{opt}$, p413-TEF-sc*TKL1*	MuA07	This Chapter

Table 2.2: continued next page.

Strain	Genotype	Plasmids	Parent Strain	Reference
MuA10	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::P_{GPD}-aro4$_{k229l}$; Ty2δ::P_{GPD}-ECL_01944$_{opt}$; zwf1Δ*	none	MuA07	This Chapter
MuA11	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::P_{GPD}-aro4$_{k229l}$; Ty2δ::P_{GPD}-ECL_01944$_{opt}$; zwf1Δ*	p425-TEF-pa5_5120$_{opt}$/GPD-ca*HQD2*$_{opt}$, p413-TEF-sc*TKL1*, p426-GPD	MuA10	This Chapter
MuA12	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::P_{GPD}-aro4$_{k229l}$; Ty2δ::P_{GPD}-ECL_01944$_{opt}$; zwf1Δ*	p425-TEF-pa5_5120$_{opt}$/GPD-ca*HQD2*$_{opt}$, p413-TEF-sc*TKL1*, p426-GPD-ECL_01944$_{opt}$	MuA10	This Chapter

Table 2.2 Yeast strains used in this chapter.

A table of yeast strains generated in this chapter and the corresponding genotype.

Reaction Name	Reaction Formula
'AROZ'	3dhsk[c] -> h2o[c] + pca[c]
'AROY'	pca[c] -> co2[c] + cat[c]
'CATA'	o2[c] + cat[c] -> mua[c]
'MUAe'	mua[c] <=> mua[e]
'EX_MUA'	mua[e] <=>

Table 2.3 Reactions added to iMM904 model to account for the heterologous muconic acid pathway.

The heterologous pathway for muconic acid was added to the standard genome scale model. The following abbreviations are used: 3dhsk is 3-dehydroshikimate, h2o is water, pca is protocatechuic acid, co2 is carbon dioxide, cat is catechol, o2 is oxygen, and mua is muconic acid. The [c] and [e] after each species denote the cytoplasm and extracellular compartments, respectively.

Species	**Gene**	**Optimization**	**K_m (mM)**	**V_{max} (mM/min/µg protein extract)**
Klebsiella pneumoniae	*AroZ*	Not optimized	n.d.	n.d.
Klebsiella pneumoniae	*AroZ*	Codon-optimized	0.65 ± 0.09	$(1.0 \pm 0.16)x10^{-4}$
Aspergillus niger	*QutC*	Codon-optimized	n.d.	n.d.
Podospora anserina	Pa_5_5120	Codon-optimized	0.30 ± 0.04	$(1.8 \pm 0.4)x10^{-4}$
Debaryomyces hansenii	DEHA2F15906g	Not optimized	n.d.	n.d.

Table 2.4 In vitro assay of DHS dehydratase genes.

A spectrophotometric assay using cell lysates was conducted to compare catalytic constants for candidate DHS dehydratase genes. K_m and V_{max} standard deviation values are based on results from biological triplicates. Enzymes without detectable activity are designated n.d.

Species	**Gene**	**Optimization**	**K_m (mM)**	**V_{max} (mM/min/µg protein extract)**
Acinetobacter baylyi	*CatA*	Not optimized	0.16 ± 0.05	$(4.1 \pm 0.7)x10^{-4}$
Acinetobacter baylyi	*CatA*	Codon-optimized	0.23 ± 0.04	$(1.2 \pm 0.3)x10^{-4}$
Debaryomyces hansenii	DEHA2C14806g	Not optimized	0.23 ± 0.06	$(2.2 \pm 0.4)x10^{-4}$
Candida albicans	*HQD2*	Codon-optimized	0.17 ± 0.03	$(2.8 \pm 0.2)x10^{-4}$

Table 2.5 In vitro assay of catechol 1,2-dioxygenase genes

A spectrophotometric assay using cell lysates was conducted to compare catalytic constants for candidate catechol 1,2-dioxygenase genes. Km and Vmax standard deviation values are based on results from biological triplicates.

Species	**Gene**	**Optimization**	**Muconic acid production (mg/L)**
Klebsiella pneumoniae	*AroY*	Not optimized	3.9 ± 0.1
Klebsiella pneumoniae	*AroY*	Codon-optimized	4.4 ± 0.8
Debaryomyces hansenii	DEHA2G00682g	Not optimized	n.d.
Podospora anserina	Pa_0_880	Codon-optimized	n.d.
Podospora anserina	Pa_4_4540	Codon-optimized	n.d.
Saccharomyces cerevisiae	*FDC1* and *PAD1*	Not optimized	n.d.
Enterobacter cloacae	ECL_01944	Codon-optimized	5.3 ± 0.3

Table 2.6 *In vivo* assay of PCA decarboxylase genes co-expressed with Pa_5_5120 from *P. anserina* and HQD2 from *C. albicans*.

Candidate PCA decarboxylase enzymes were assayed through an *in vivo* pathway complementation assay. Muconic acid production standard deviation values are based on results from biological triplicates. Enzymes without detectable activity are designated n.d.

Product	**Precursor metabolite**	**Concentration (mg/L)**	**Maximum Yield**	**Reference**
Muconic acid	Dehydroshikimate	141±8	3.9±0.2 mg/$g_{Glucose}$	This chapter
Vanillin	Dehydroshikimate	45	2.3 mg/$g_{Glucose}$	[29]
p-hydroxybenzoic acid	Tyrosine	89.8	6.0 mg/$g_{Glucose}$	[30]
p-amino benzoic acid	Chorismate	34.3	2.3 mg/$g_{Glucose}$	[30]
p-hydroxycinnamic acid	Phenylalanine	33.3	1.7 mg/ $g_{Raffinose}$	[31]
Resveratrol	Phenylalanine	14.4	0.22 mg/$g_{Coumaric\ acid}$	[32]
Naringenin	Tyrosine	7	0.35 mg/$g_{Galactose}$	[33]
Indolylglucosinolate	Tryptophan	1.07±0.38	0.054±0.02 mg/$g_{Glucose}$	[34]

Table 2.7 Compilation of shikimate or aromatic amino acid-based metabolite production in yeast for simple shake-flask conditions.

Metabolite titers and yields are compared for similar molecules produced in metabolically engineering *S. cerevisiae*. The strain described in this chapter has the highest titer and the second highest yield to date for this class of molecules.

Chapter 3: Coordinated Transcription Factor and Promoter Engineering to Establish Strong Expression Elements in *Saccharomyces cerevisiae*[3]

3.1 CHAPTER SUMMARY

In pursuit of a biosensor which can facilitate further improvements in muconic acid production, we identified a transcription factor and promoter sensitive to the downstream products of the shikimate pathway and used them to develop tools for gene expression. In this chapter, we present a coordinated approach that combines cis-acting element engineering with mutant trans-acting factors to engineer yeast promoters. Specifically, we first construct a hybrid promoter based on the *ARO9* upstream region that exhibits high constitutive and inducible expression with respect to exogenous tryptophan. Next, we perform protein engineering to identify a mutant Aro80p that affords both high constitutive expression while retaining inducible traits. We then use this mutant trans-acting factor to drive expression and generate ultra-strong promoters with transcriptional output roughly 2 fold higher than *TDH3* (GPD), one of the strongest promoters to-date. Finally, we used this element to construct a modular expression system capable of staged outputs resulting in a system with nearly 6-fold, 12-fold and 15-fold expression relative to the off-state. This chapter further highlights the potential of using endogenous transcription factors (including mutant factors) along with hybrid promoters to expand the yeast synthetic biology toolbox.

[3] Leavitt, J. M., Tong, A., Tong, J., Pattie, J., Alper, H. S., Coordinated transcription factor and promoter engineering to establish strong expression elements in *Saccharomyces cerevisiae*. *Biotechnology journal* 2016. The author made significant contributions to designing, conducting and analyzing the experiments as well as preparing and editing the manuscript.

3.2 INTRODUCTION

Precise control of gene expression is essential for many applications including protein production, metabolic engineering, and synthetic biology [124]. This control is the result of both cis-acting elements and trans-acting factors that interact to determine transcriptional capacity. In recent years, our capacity to influence these elements and factors has increased, thus yielding an increase in the number of distinct promoter modalities available for many model host organisms. However, the common eukaryotic model yeast *Saccharomyces cerevisiae* is still transcriptionally limited compared with the tools available in common bacterial hosts such as *Escherichia coli* [58, 125-129]. Nevertheless, the attractiveness of yeast as a platform organism (especially for metabolic engineering applications) [4, 5, 7] warrant continued efforts to expand the set of tools available. Most synthetic biology efforts in *S. cerevisiae* have focused primarily focused on cis-elements as a means of influencing transcriptional output via modifications to the promoter sequence and control of RNA half-life through terminator sequence [62, 130, 131]. In this vein, significant efforts have been made in cataloging the binding sites for various transcription factors within promoter sequences as well as characterizing their relative importance [132, 133]. While these efforts have been effective at modulating (and sometimes, increasing) promoter activity, there are still only a limited number of orthogonal systems in yeast as well as a limited set of inducible promoters.

The engineering of trans-acting factors enables an additional level of expression control that is compatible with generating more inducible promoters and further complex genetic circuits. Indeed, previous efforts for trans-acting factor engineering in yeast has focused on importing orthogonal trans-elements from other systems such as TALE's, dCas9, bacterial response elements [134], and Estrogen Receptors [70, 135, 136]. When

coupled with the proper cis-factor architecture, these systems have provided additional levels of inducible control and have facilitated multi-gene activation. However, with few exceptions, these systems have not provided tools which facilitate novel, strong transcriptional activation above currently existing systems.

Despite promise of orthogonal trans-factors for expanding the range of inducible systems in yeast, the set of inducible promoters for *S. cerevisiae* remains limited [50, 137]. Most of these systems rely on small molecule detection, such as copper, phosphates, methionine or aromatic amino acids [48, 138-140] using native, trans-activing factors that result in relatively weak levels of inducibility. The stand-out anomaly both with respect to carbon source inducible nature and strength of induction is the galactose system [141]. Another inducible system of interest is the ethanol inducible TPS1 system, which has been used to induce flocculation in the industrial yeast strain ZLH01 and achieve cost-effective cell separation [142]. While it is possible to use these endogenous systems for unique applications including quorum sensing and metabolic control [143], there is still a lack of parts to enable complex synthetic circuits in yeast.

One particularly attractive and potent endogenous inducible system in *S. cerevisiae* that has been studied throughout the years is aromatic amino acid induction and regulation [48, 144, 145]. The first step in aromatic amino acid catabolism, the aromatic amino acid transferase II protein encoded by *ARO9,* is under significant regulation. In particular, the *ARO9* (as well as *ARO10)* gene is activated by GATA and Aro80p transcription factors in response to nitrogen limitation as well as exogenous aromatic amino acids [144]. This activating function in conjunction with knowledge of

the Aro80p binding site has been recently used in a synthetic context to induce gene expression [146]. However, these systems have not been fully engineered for maximal expression and inducibility through trans-factor engineering. Here, we present the engineering of the Aro80p trans-activing factor for enhanced expression and inducibility. We demonstrate that an identified mutant factor can be used in conjunction with cis-acting element engineering to create ultra-strong promoters with activity nearly 2-fold higher than the strong, constitutive *TDH3* (GPD) promoter. Finally, we demonstrate the capacity to utilize this factor in a configuration that is capable of generating staged outputs with up to 6-fold, 12-fold or 14-fold induction from the "off" state as a function of inputs. Collectively, these results demonstrate the capacity to rewire yeast promoters through the modulation of both cis-acting element and trans-acting factor components.

3.3 Results and Discussion

3.3.1 Initial synthetic promoter construction using an aromatic inducible transcription factor

Initially, we sought out to engineer both minimally-sufficient and hybrid cis-elements to establish a baseline aromatic amino acid inducible promoter system for yeast. The design for this promoter element was based on the native *ARO9* promoter that has been previously demonstrated to have aromatic amino acid (esp. tryptophan) inducibiliy [48]. First, to generate a native design, a 355 base pair promoter was amplified from the genome corresponding to the previously reported structure with the additional truncation of four potentially confounding Pho4p binding sites [147, 148]. This native promoter

was cloned into a basic yeast vector in front of the yECitrine fluorescent protein (in a low copy centromeric p416 plasmid to form the "ARO9wt-YFP" construct) [122]. This wild-type promoter construct showed slight constitutive expression in basic minimal media and over 2-fold induction upon exposure to 500 mg/L tryptophan (**Figure 3.1A**). Thus, this endogenous element could serve as a baseline for aromatic acid inducibility.

Next, to enable a more minimal and hybrid approach, we dissected the *ARO9* promoter to extract a more minimally sufficient UAS_{aro} element. The UAS_{aro} element previously described [48] failed to provide any activity when cloned upstream of a LeuMin core, therefore this putative UAS_{aro} element was expanded to include 56 bp 5' and 24 bp 3' of the Aro80p binding site. The 5' flanking region begins after a putative URS1 element, while the 3' flanking region ends before a TATA box like sequence. When this enlarged UAS element was linked with the LeuMin core promoter, we designed a functional 249 bp synthetic promoter. The resulting construct was cloned into a similar plasmid as described above to form the "1x UAS_{aro} -YFP" construct. In similar fashion to the endogenous promoter, transformed yeast cells were evaluated by flow cytometry at mid-with and without a spiked of 500mg/L L-tryptophan to demonstrate Aro80p based activation. Both the native and the synthetic promoter demonstrated leaky constitutive expression (likely due to endogenous aromatic amino acid levels) but had inducible characteristics upon exposure to culture media spiked with 500 mg/L L-tryptophan (**Figure 3.1A**). The basal expression of the hybrid promoter construct was significantly higher than that of the native promoter sequence. This difference could be

due to additional repressor binding sites not being captured within the isolated UASaro element. The tight regulation of the native *ARO9* promoter is expected as this element would be under strong evolutionary pressure to prevent constitutive catabolism of aromatic amino acids leading to a futile cycle.

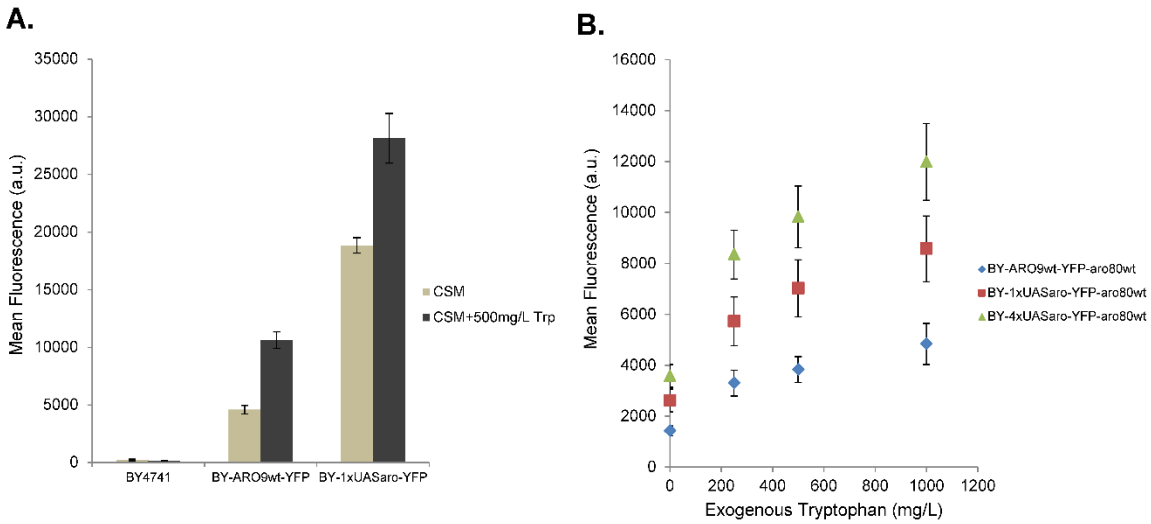

Figure 3.1 Developing a tryptophan sensitive hybrid promoter.

(A) The 355bp "wild-type" ARO9 promoter and 1xUASaro hybrid promoters are evaluated by flow cytometry following subculture into CSM and CSM containing 500mg/L of tryptophan. Both promoters demonstrate leaky, constitutive expression with inducible traits with upward of 2-fold induction upon tryptophan addition. The 1xUASaro hybrid promoter exhibited higher constitutive expression while maintaining modest induction capacity. (B) Promoter constructs (synthetic designs with 1 and 4x UASaro elements and the endogenous control) were tested alongside a plasmid containing the aro80wt gene under control of the strong TEF promoter. Each construct maintained a roughly 2.5 fold inducible range while hybrid engineering increased the net expression from the constructs. The BY4741- plasmid control is included to demonstrate the relative level of background auto-fluorescence and samples for (A) and (B) were analyzed by flow cytometry on the separate days under the same conditions. Error bars represent standard deviations across biological replicates.

3.3.2 Hybrid promoter engineering to refine the aromatic amino acid response

We have previously demonstrated that a hybrid promote reengineering approach can increase the overall activity of promoters through the modification of cis-acting elements [62, 66]. In this regard, we sought to demonstrate that additional copies of the above identified UAS_{aro} can amplify the transcriptional output of the synthetic promoter. To do this, we increased the number of copies of UAS_{aro} elements from 1x to 4x upstream of the core promoter. Furthermore, as endogenous Aro80p pools might limit the transcriptional output from these hybrid promoters, we overexpressed the native *ARO80* gene in trans along with these constructs. The resulting inducible capacity of these three systems (the wild-type *ARO9* promoter and the 1x and 4x hybrid constructs) was evaluated via flow cytometry using a range of tryptophan concentrations (**Figure 3.1B**). These results demonstrate that indeed a hybrid promoter engineering approach can amplify the transcriptional output of these promoters. In each of the cases, a 2.5-fold inducibility level was observed with and without exogenous tryptophan whereas the magnitude of expression followed the trend of *ARO9* < 1x UAS_{aro} < 4x UAS_{aro}. As a result, it is possible to tune the output of this inducible promoter system via a hybrid promoter engineering approach.

3.3.3 Establishing a mutant Aro80p factor that can alter promoter response

Next, we sought to modulate the trans-acting factor for this promoter element (namely Aro80p) as a means of altering promoter activity. Specifically, the goal was to modulate both the constitutive and inducible traits of this promoter via protein engineering. More specifically, we wished to retain some inducible traits in the promoter, but at the same time, enabling a much stronger constitutive response. To accomplish this, we used error-prone PCR to generate a mutant library of *aro80* genes. To screen for altered function, we utilized an *aro80Δ* cell line that avoids interaction with native *ARO80* expression and function. The *aro80* mutant library was transformed into the *aro80Δ* cell line expressing p416-4xLeuMin-YFP and enriched via FACS (FACSaria) to identified improved mutants. In this case, we opted to sort the top 1% of fluorescent cells in the presence of D-tryptophan to block activation domains. Out of many colonies evaluated, one particular set of mutations (specifically I551S and S675T) was isolated from a variant that yielded high constitutive and inducible activation of the 4x UAS_{aro} promoter (**Figure 3.2**). The particular variant, $aro80_{I551S,S675T}$ is hereafter referred to as "aro80mut".

To identify the causative mutation(s) responsible for aro80mut function, we individually reverted the two point mutations and evaluated function of the resulting single mutation transcription factors. **Figure 3.2** demonstrates that the dominant mutation responsible for the function of aro80mut was I551S. Specifically, reverting the mutation at residue 675 still retained function whereas reversion of the mutation at residue 551 reverted functions to levels near that of the wild-type Aro80p. Collectively,

we have demonstrated the capability of identifying a mutant version of Aro80p capable of increasing constitutive expression levels by over 5-fold compared to the wild-type version, while maintaining an inducible response to exogenous amino acids. It is possible to use this new mutant in conjunction with cis-acting element promoter engineering to develop new synthetic circuits as described in the sections below.

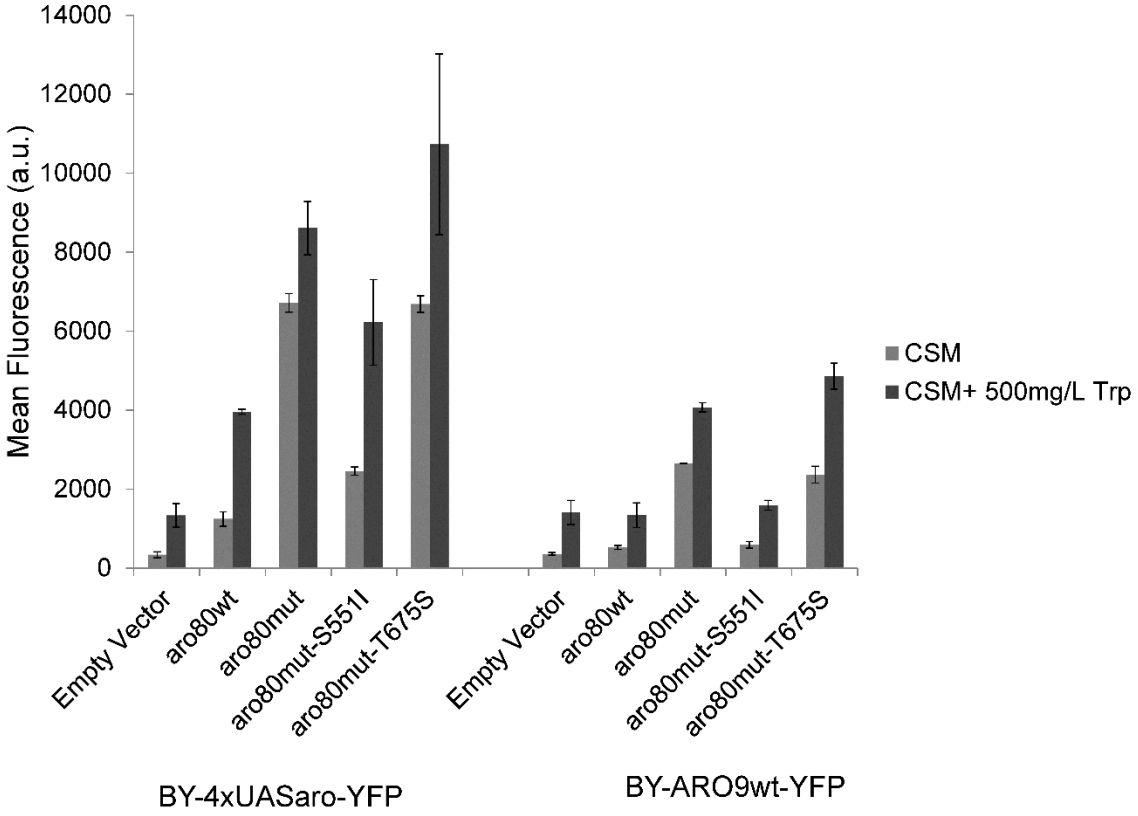

Figure 3.2 Isolating causative mutations in the aro80 mutant.

A single residue reversion assay was conducted to identify the causative mutation(s) leading to aro80 function. The aro80mut, two revertants, ARO80 wt and a blank plasmid were expressed in conjunction with the 4xUASaro-YFP and ARO9wt-YFP plasmids and measured in CSM and CSM containing 500mg/L of tryptophan. All samples were analyzed by flow cytometry on the same day under the same conditions. Error bars represent standard deviations across biological replicates.

3.3.4 Development of an ultra-strong promoter via aro80mut

To demonstrate the power of coupling cis-acting element and trans-acting factor engineering to alter promoter function, we first use this factor to drive high level, constitutive gene expression. To do so, we establish a simple "amplifier" by expressing the *aro80mut* gene under the control of a strong constitutive promoter (in this case, the *TEF* promoter) and using the resulting protein to drive the expression of a hybrid promoter (**Figure 3.3A**). Previous work has shown that the addition of UAS elements upstream of native promoters can significantly increase their net transcriptional output [62, 149]. The rationale being that the promoters are limited by transcription factor binding and that by increasing the presence of the enhancer elements, we increase the frequency of binding and therefore transcription. In this case, we use our hybrid promoter engineering approach to place several UAS_{aro} elements (in this case, 4 to 5 copies) upstream of several full-length, endogenous promoters (CYC1 and HXT7) a minimal promoter (LeuMin), and the synthetic, minimal core promoter CORE1 **40**. In each of these cases, constitutive expression of the mutant *aro80* factor greatly increased the net, constitutive expression from each of the promoters (**Figure 3.4A**). For the case of our minimal core promoter, this amplification gain was up to 15-fold. In several of the cases, the obtained expression in this circuit was greater than that of the *TDH3* (GPD) promoter, arguably among the strongest, endogenous constitutive promoters for yeast. The strongest resulting hybrid promoter (the 5x- UAS_{aro} hybrid) exhibited up to 2-fold increase in yECitrine fluorescence compared with the TDH3 promoter and upwards of 1.7-fold increase in mRNA output (**Figure 3.4B**). Thus, these experiments demonstrate

that using an aro80mut can result in strong, constitutive expression of an output promoter leading to one of the strongest promoters available for *S. cerevisiae*.

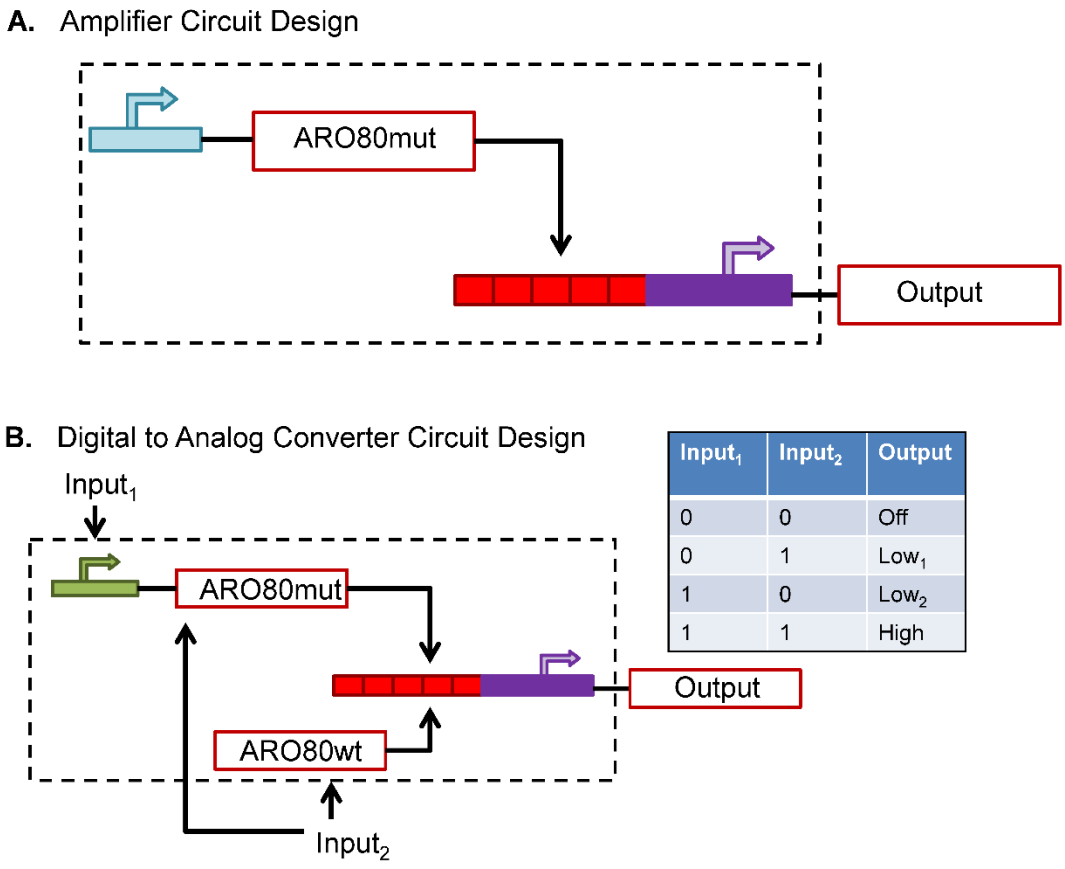

Figure 3.3 Synthetic circuit schematics.

Schematics for the two circuits considered in this work are provided. (A) An amplifier was created through the use of the constitutive overexpression of aro80mut driving the activation of a hybrid promoter. (B) A Digital-to-Analog converter was created consisting of aro80mut under control of the GAL1 promoter driving activation of the hybrid promoters. With this system, the output is modulated by two inputs (carbon source: glucose or galactose and exogenous tryptophan: low and high). This expected output is provided in a representative truth table.

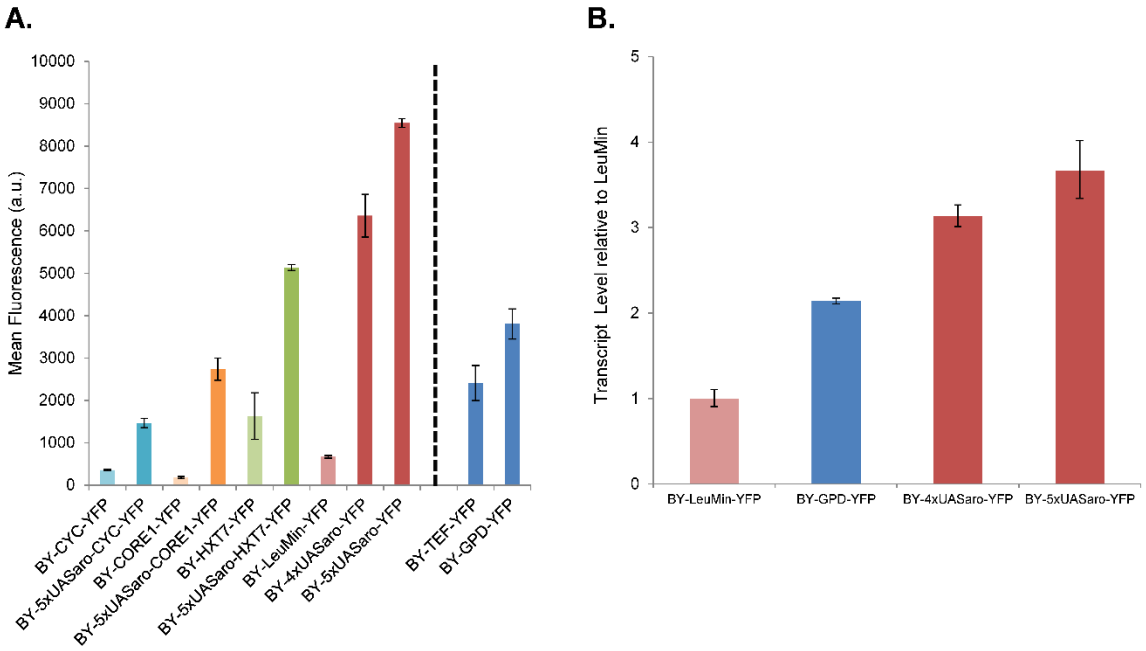

Figure 3.4 Development of ultra-strong promoters via aro80mut.

Using hybrid promoter engineering, tandem repeats of the UASaro were placed upstream of a variety of promoters. (A) These plasmids were expressed within the aro80mut amplifier context presented in Figure 3A resulting in gene expression up to 15 fold higher with the hybrid constructs compared to the corresponding basal promoter (shown in lighter hues). Some promoters exhibited up to 2-fold higher expression than GPD, one of the strongest promoters in yeast. All samples were analyzed by flow cytometry on the same day under the same conditions. Error bars represent standard deviations across biological replicates. (B) Transcript levels were measured for three strains expressing amplifier circuits along with GPD-YFP as a benchmark. All transcript values are reported relative to the BY-LeuMin-YFP-aro80mut strain. The expression profiles match fluorescence data, with the 5xUASaro amplifier expression being roughly 1.7 fold that of the GPD promoter. Error bars represent standard deviation values based on three technical replicates.

3.3.5 Development of a promoter with staged output using the aro80mut

As a second demonstration of the utility of coupling cis- and trans-acting factor engineering, we utilized the resulting aro80mut to achieve a staged, multi-output response in yeast. A similar circuit has recently been demonstrated for *E. coli* whereby two digital (on/off) inputs result in an analog response, termed a "Digital-to-Analog converter" [150]. While the biological circuit is inherently noisier than its electronic counterpart, this term is employed here to be consistent with the literature [150]. To demonstrate a similar circuit for yeast, we expressed the *aro80mut* gene under transcriptional control of the *GAL1* promoter (**Figure 3.3B**). This mutant factor was then used to drive the expression of three distinct hybrid promoter constructs. Using this system, it is possible to modulate output via changes to the conditions of glucose vs. galactose and un-spiked vs spiked tryptophan. In this regard, the glucose condition represents an "off" state, the tryptophan induction of endogenous, wild-type Aro80p represents a "Low_1" intermediate state, galactose induction of aro80mut represents a "Low_2" intermediate state and the tryptophan induction of both endogenous aro80wt and galactose induced aro80mut presents the final "High" state. To test this function, the four experimental inputs were: glucose, glucose with 1g/L tryptophan, galactose, and galactose with 1g/L tryptophan. **Figure 3.5** demonstrates the realization of this 4-state promoter system. For the case of the 5x-LeuMin, the fold expression over the "off" state was 6 fold higher with tryptophan, 12 fold with galactose and 14.6 fold with galactose and tryptophan. This full induction resulted in fluorescence values roughly 50% of the

GAL1 promoter. The 5xCYC circuit responded comparably with 3 fold, 6 fold and almost 8 fold activation for each of the various on states.

Finally, by removing the activation provided by endogenous, wild-type Aro80p, it is possible to remove a potential state. Specifically, by expressing this system in an *aro80* deletion strain (**Figure 3.5**), the system can no longer be activated solely by tryptophan and thus the only states observed are "Off", "Low_2" and "High". Collectively, these two applications demonstrate the utility of using a mutant trans-acting factor.

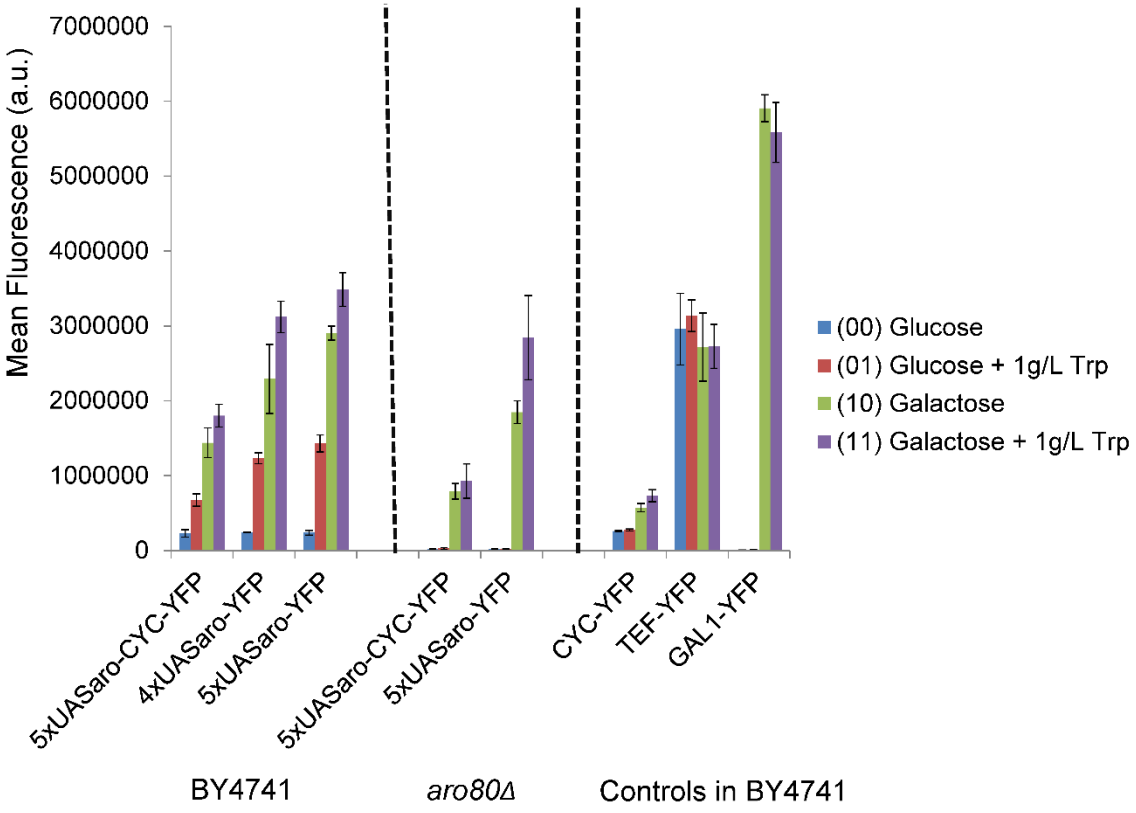

Figure 3.5 Demonstrating a staged-output promoter system.

A so-called "Digital-to-Analog converter circuit" (Figure 3B) was generated with the aro80mut gene under transcriptional control of the GAL1 promoter and used to drive the expression of three hybrid promoter constructs. Resulting expression of the 5x-UASaro circuit is 6 fold higher with tryptophan over the "off" state due to the activation of the endogenous ARO80p, 12 fold higher with galactose inducing expression of aro80mut and 14.6 fold with galactose and tryptophan induced. The 5xUASaro-CYC construct responded comparably with 3 fold, 6 fold and 8 fold activation. These outputs compare with the truth table depicted in Figure 3B. These constructs are expressed in an aro80Δ background to remove one of the states. CYC-YFP and TEF-YFP are included in the circuit context as endogenous controls and GAL1-YFP is expressed without the circuit as an expression benchmark. All samples were analyzed by flow cytometry on the same day under the same conditions. Error bars represent standard deviations across biological replicates.

3.4 CONCLUDING REMARKS

In this chapter, we demonstrate the power of combining cis-acting element engineering with mutant trans-acting factors to engineer yeast promoters. Specifically, we first develop an inducible, hybrid promoter based on the upstream region of the *ARO9* promoter. Next, we isolate a mutant *aro80* protein that can afford increased constitutive expression while retaining inducible traits. Finally, we utilize this factor to generate ultra-strong promoters and to establish a promoter capable of staged-outputs. In the former case, we demonstrate promoters with transcriptional output roughly 2-fold higher (based on both fluorescence and mRNA) compared to the *TDH3* (GPD) promoter. Thus, this promoter system is one of the strongest yeast expression systems reported to date. In the latter case, we enabled a system with activation levels of 6-fold, 12-fold or 14-fold of the "off" state as a function of circuit input. Collectively, this chapter demonstrates the utility of both engineering endogenous transcription factors with hybrid promoter engineering approaches. The ability to expand this paradigm for other endogenous or previously demonstrated heterologous systems provides great promise for expanding the yeast synthetic biology toolbox.

Plasmid	**Source or Assembly Method**	**Primers**
p416-ARO9wt-yECitrine	Restriction Cloning with PmeI/XbaI	Fwd: aaagctGTTTAAACtgaacatggttatgttatat attgtttg
		Rev: GCTCTAGAtgagtcgatgagagagtgtaatt
p416-LeuMin-yECitrine	[62]	n/a
p416-1xUASaro-LeuMin-yECitrine	Restriction Cloning with HindIII/PmeI	Fwd: cccAAGCTTcggccgtagataataacaaag
		Rev: aaagctGTTTAAACatgtttcctaccccaatga t
p416-4xUASaro-LeuMin-yECitrine	Sequential Restriction Cloning: PacI/HindIII	Fwd: ccTTAATTAAcggccgtagataataacaaag
		Rev: cccAAGCTTatgtttcctaccccaatgat
	Sequential Restriction Cloning: AscI/PacI	Fwd: ttGGCGCGCCcggccgtagataataacaaag
		Rev: ccTTAATTAAatgtttcctaccccaatgat
	Sequential Restriction Cloning: BamHI/AscI	Fwd: CGCGGATCCcggccgtagataataacaaag
		Rev: ttGGCGCGCCatgtttcctaccccaatgat

Table 3.1: continued next page.

p415-TEF	ATCC	n/a
p415-TEF-aro80	Homlogous Recombination	aro80 Fwd: gctcattagaaagaaagcatagcaatctaatctaagtt tATGTCTGCTAAGAAAAGGCC
		aro80 Rev: ggcgtgaatgtaagcgtgacataactaatTTATT TACGCGTTATTGGCC
		Vector Rev: AAACTTAGATTAGATTCGTATGC TTTCTTTC
		Vector Fwd: ATTAGTTATGTCACGCTTACATT CACG
p415-TEF-aro80-mut	GeneMorph II amplification and ligation	Mutagenesis Fwd: gACTAGTATGTCTGCTAAGAAAA GGCC
		Mutagenesis Rev: tgaatgtaagcgtgacataactaatctcgagTTA
p415-TEF-aro80-mut-S551I	Quick Change Kit	Fwd: GCAAAAATAGAGATCATTCGAA TCCT

Table 3.1: continued next page.

		Rev: AGGATTCGAATGATCTCTATTTTTGC
p415-TEF-aro80-mut-T675S	Inverse PCR followed by blunt end ligation	Fwd: TCTGCAAAAGAAATATTGAGTTC
		Rev: TCTGTATGCTAATTCCACATAC
p416-CYC-yECitrine	[62]	n/a
p416-HXT7-yECitrine	[130]	n/a
p416-GPD-yECitrine	[62]	n/a
p416-TEF-yECitrine	[64]	n/a
p416-CORE1-yECitrine	Inverse PCR followed by blunt end ligation	Fwd: GGCGCCGGAAAAAAGCATCGAAAAAAtctagaatgtctaaaggtgaagaattattcactg
		Rev: TCCACTCACGCCCAACAGTGCTCTTTTATAGAGCTCCAGCTTTTGTTCC
p416-5xUASaro-CYC1-yECitrine	Gibson Assembly	5xUASaro Fwd: cctcactaaagggaacaaaagctggagctcGGATCCCGGCCGTAGATAATAAC

Table 3.1: continued next page.

		5xUASaro Rev: gcttgatccaccaaccaacgctcgccaaatGTTTAAACATGTTTCCTACCCCAATGATG
		Vector Rev: ttcctttgttattatctacggccgGGATCCGAGCTCCAGCTTTTGTTCC
		Vector Fwd: ccatcattggggtaggaaacatGTTTAAACATTTGGCGAGCGTTGGTT
p416-5xUASaro-HXT7-yECitrine	Restriction Cloning with BamHI/PmeI	n/a
p416-5xUASaro-LeuMin-yECitrine	Restriction Cloning with BamHI/PmeI	n/a
p416-5xUASaro-CORE1-yECitrine	Restriction Cloning with BamHI/PmeI	n/a
p415-Gal1-aro80-mut	Gibson Assembly	pGAL1 Fwd: cctcactaaagggaacaaaagctggagctcAGTACGGATTAGAAGCCGCCG
		pGAL1 Rev: GAAGGCCTTTTCTTAGCAGACATactagtGTTTTTTCTCCTTGACGTTAAAGTATAGAGG

Table 3.1: continued next page.

		Vector Rev: TCGCCCGCTCGGCGGCTTCTAAT CCGTACTGAGCTCCAGCTTTTGT TCCCTTTA
		Vector Fwd: ACTTTAACGTCAAGGAGAAAAA ACactagtATGTCTGCTAAGAAAAG GCCTTCG
p416-GAL1-yECitrine	[62]	n/a

Table 3.1 Plasmids used in this chapter.

A list of plasmids generated and used in this chapter.

Strain Name	**Plasmid(s)**
BY4741	N/A
BY-ARO9wt-YFP	p416-ARO9wt-yECitrine
BY-1xUASaro-YFP	p416-1xUASaro-LeuMin-yECitrine
BY-ARO9wt-YFP-aro80wt	p416-ARO9wt-yECitrine , p415-TEF-aro80
BY-1xUASaro-YFP-aro80wt	p416-1xUASaro-LeuMin-yECitrine , p415-TEF-aro80
BY-4xUASaro-YFP-aro80wt	p416-4xUASaro-LeuMin-yECitrine , p415-TEF-aro80
BY4741Δaro80	N/A
aro80Δ-4xUASaro-YFP	p416-4xUASaro-LeuMin-yECitrine
BY-4xUASaro-YFP-Empty Vector	p416-4xUASaro-LeuMin-yECitrine , p415-TEF
BY-4xUASaro-YFP-aro80wt	p416-4xUASaro-LeuMin-yECitrine, p415-TEF-aro80
BY-4xUASaro-YFP-aro80mut	p416-4xUASaro-LeuMin-yECitrine, p415-TEF-aro80-mut
BY-4xUASaro-YFP-aro80mut-S551I	p416-4xUASaro-LeuMin-yECitrine , p415-TEF-aro80-mut-S551I
BY-4xUASaro-YFP-aro80mut-T675S	p416-4xUASaro-LeuMin-yECitrine , p415-TEF-aro80-mut-T675S
BY-ARO9wt-YFP-Empty Vector	p416-ARO9wt-yECitrine , p415-TEF
BY-ARO9wt-YFP-aro80wt	p416-ARO9wt-yECitrine , p415-TEF-aro80
BY-ARO9wt-YFP-aro80mut	p416-ARO9wt-yECitrine , p415-TEF-aro80-mut
BY-ARO9wt-YFP-aro80mut-S551I	p416-ARO9wt-yECitrine , p415-TEF-aro80-mut-S551I
BY-ARO9wt-YFP-aro80mut-T675S	p416-ARO9wt-yECitrine , p415-TEF-aro80-mut-T675S
BY-CYC-YFP-aro80mut	p416-CYC-yECitrine, p415-TEF-aro80-mut
BY-5xUASaro-CYC-YFP-aro80mut	p416-5xUASaro-CYC1-yECitrine, p415-TEF-aro80-mut

Table 3.2: continued next page.

BY-CORE1-YFP-aro80mut	p416-5xUASaro-CYC1-yECitrine, p415-TEF-aro80-mut
BY-5xUASaro-CORE1-YFP-aro80mut	p416-5xUASaro-CYC1-yECitrine, p415-TEF-aro80-mut
BY-HXT7-YFP-aro80mut	p416-HXT7-yECitrine , p415-TEF-aro80-mut
BY-5xUASaro-HXT7-YFP-aro80mut	p416-5xUASaro-HXT7-yECitrine , p415-TEF-aro80-mut
BY-LeuMin-YFP-aro80mut	p416-LeuMin-yECitrine , p415-TEF-aro80-mut
BY-4xUASaro-YFP-aro80mut	p416-4xUASaro-LeuMin-yECitrine , p415-TEF-aro80-mut
BY-5xUASaro-YFP-aro80mut	p416-5xUASaro-LeuMin-yECitrine, p415-TEF-aro80-mut
BY-TEF-YFP-aro80mut	p416-TEF-yECitrine , p415-TEF-aro80-mut
BY-GPD-YFP	p416-GPD-yECitrine
BY-5xUASaro-CYC-YFP-GAL1-aro80mut	p416-5xUASaro-CYC1-yECitrine, p415-Gal1-aro80-mut
BY-4xUASaro-YFP-GAL1-aro80mut	p416-4xUASaro-LeuMin-yECitrine,p415-Gal1-aro80-mut
BY-5xUASaro-YFP-GAL1-aro80mut	p416-5xUASaro-LeuMin-yECitrine, p415-Gal1-aro80-mut
aro80Δ-5xUASaro-CYC-YFP-GAL1-aro80mut	p416-5xUASaro-CYC1-yECitrine, p415-Gal1-aro80-mut
aro80Δ-5xUASaro-YFP-GAL1-aro80mut	p416-5xUASaro-LeuMin-yECitrine, p415-Gal1-aro80-mut
BY-CYC-YFP-GAL1-aro80mut	p416-CYC-yECitrine, p415-Gal1-aro80-mut
BY-TEF-YFP-GAL1-aro80mut	p416-TEF-yECitrine ,p415-Gal1-aro80-mut
BY-GAL1-YFP	p416-GAL1-yECitrine

Table 3.2 Yeast strains used in this chapter.

Yeast strains used in this chapter and the plasmids which they maintain.

Chapter 4: Biosensor Directed Evolution for Muconic Acid Production in *Saccharomyces cerevisiae*

4.1 Chapter Summary

Having developed the ARO9 cis-acting elements and ARO80 trans-acting factor components to facilitate improvements in gene expression, we next employed this as a biosensor to further develop muconic acid production in *S. cerevisiae.* To further increase muconic acid production in this host with industrially relevant titers, we employed an adaptive laboratory evolution (ALE) strategy to complement rational metabolic engineering. ALE allows for the selection of global phenotypes without prior knowledge of an organism's metabolism. Isolating improved strains relies on the availability of an effective selection strategy specific for the desired phenotype. In this chapter, we adapted a biosensor device developed in the previous chapter which is sensitive to the endogenous aromatic amino acid production (AAA) in *S. cerevisiae* using a hybrid promoter approach, and used this biosensor to augment an anti-metabolite ALE scheme. Following two iterations of mutation and selection in our ALE scheme, we isolated strains of *S. cerevisiae* that are capable of 2-fold AAA production relative to our previously engineered strain and 10-fold that of wild-type *S. cerevisiae*. Having successfully selected for improvements in flux through the shikimate pathway, we then demonstrate that this can be redirected into the composite pathway and on to muconic acid formation.

The resulting four strains represent a combination of rational metabolic engineering and evolutionary adaptation. To demonstrate the improvements in flux gained from ALE, we expressed the composite muconic acid pathway we have previously demonstrated, and the ALE isolated strains were capable of three fold the pathway production compared to our previously engineered strain. We next desired to rationally reroute flux into the composite pathway through a truncation of the shikimate pathway. We truncated the penta-functional ARO1 protein and expressed it resulting in strains capable of 7.5 fold output from the composite pathway. Our final step in strain engineering was the expression of the endogenous PCA decarboxylase, scPAD1, which resulted in a strain capable of over 550mg/L muconic acid production in flasks and 1.94 g/L in a fed-batch bioreactor. This represents the highest production of muconic acid in *S. cerevisiae* to date in addition to the highest reported titer of a shikimate pathway derivative.

4.2 INTRODUCTION

In chapter 3, we employed a hybrid approach to develop promoters which were inducible to AAA. Now, we will employ the ARO9 derived hybrid promoters as biosensors to further improve *S. cerevisiae* strains for muconic acid production through ALE. In chapter 2, the production of muconic acid in *S. cerevisiae* resulted in titers of 141 mg/L [1] and represented traditional metabolic engineering work of composite pathway development, screening heterologous enzymes for function in the desired

production host, and optimizing carbon flux into the composite pathway utilizing flux balance analysis. Other groups have demonstrated further improvements in titer resulting from protein engineering which eliminated flux downstream of the composite pathway and subsequent bioreactor scale-up facilitating dissolved oxygen (DO) control to improve enzymatic activity of the rate limiting step, resulting in titers of 559.5 mg/L [24]. Other work by Horwitz and coworkers demonstrated Cas9 mediated engineering to facilitate alternative shikimate pathway utilization by importing the heterologous pathway from *E. coli*, ostensibly for the production of muconic acid, however they did not report titers of muconic acid from glucose [17]. While there has been significant interest in the development of *S. cerevisiae* for the production of secondary metabolites derived from the shikimate pathway, it faces a number of difficulties including low precursor availability and high degrees of flux control exhibited by pathway enzymes and regulatory proteins [100, 151, 152]. These difficulties have previously limited titers from glucose to the level of 102 mg/L for naringenin [153], 141 and 559.5mg/L for muconic acid [1, 24] and only through the replacement of yeast metabolism with parts from *E. coli* were groups able to increase titers of to 1.93g/L for p-coumaric acid [152]. To address these limitations, we sought to employ an adaptive evolution approach in order to improve titers beyond current limits.

The process of natural evolution leads to the selection of beneficial traits over long periods of time as organisms adapt to their surrounding environment [154]. Two of the defining characteristics of an ALE scheme are the generation of sequence diversity

and the selection mechanism. Sequence diversity can be achieved through chemical mutagenesis or through study over long evolutionary time spans while selection of the desired phenotype is typically achieved by selection under specific growth conditions leading to an enrichment of a subpopulation capable of improved growth rates [22, 39]. While improved growth rates continue to be the core mechanism of selection, recent evolutionary strategies exist which exploit a chemical feature of the desired product to facilitate selection for improved production of a compound of interest [12, 155] or the presence of an anti-metabolite to aid selection for improvements in production of a specific product [156, 157]. These anti-metabolites present a disruptive metabolic pressure on the cells which can be overcome through mutations resulting in higher concentrations of the metabolite of interest. As whole cell biosensors are becoming more commonly available [21, 158], these resources provide an opportunity to expand the scope of selectable phenotypes for strain improvement through ALE.

Biosensors allow for tying intracellular product formation with an output that can be readily screened, providing a way of screening beneficial mutations either in high throughput assays with a reporter or selection through resistance. Recently, a fluorescent biosensor was utilized in an ALE scheme to improve L-Valine production in *C. glutamicum* [44] with flow cytometry used to facilitate selection. One of the advantages of traditional ALE is the use of growth based phenotypes to allow selection rather than relying on the limited throughput of screening, which a biosensor be cleverly implemented to provide. Previously, the growth coupled screening using a biosensor in *S.*

cerevisiae was performed using the *glmS* riboswitch to screen enzymes to produce N-acetyl glucosamine [43] and L-lysine riboswitch in *E. coli* [159]. This chapter represents the first application of a transcription factor based biosensor for ALE in *S. cerevisiae* and the first coupling of a biosensor with an anti-metabolite strategy. A major advantage of this biosensor directed evolution scheme is that it provides a generalizable strategy opening up new chemical products for growth based selection.

One implementation of a biosensor is the ARO9 promoter, previously used to establish hybrid promoters of varying strength and inducibility in chapter 3. In this chapter, we demonstrate how biosensor-mediated ALE combined with local pathway optimization can further muconic acid production in *S. cerevisiae*. We demonstrate the usage of an ARO9 based biosensor and application for selection in an ALE experiment in order to screen for genome wide changes that can facilitate improvements AAA as a surrogate for improvements in muconic acid production. Through this ALE, we developed strains capable of 2-fold higher AAA production, 3-fold higher output from the muconic acid composite pathway and following some pathway engineering and bioreactor scale up resulted in 1.94 g/L muconic acid production which is the highest production of muconic acid in yeast as well as one of the highest of a shikimate pathway derivative.

4.3 Results and Discussion

4.3.1 Adaptation of Biosensor for Adaptive Laboratory Evolution

Previous approaches have focused on only rational engineering targets to improve strains for production of shikimate derivatives since a high-throughput detection was limiting. To enable high-throughput strain engineering (such as through ALE), we required a methodology to screen or select based on muconic acid level. For this chapter, we hypothesized that AAA production could serve as a surrogate to muconic acid. In the prior chapter, we built hybrid promoters through tandem insertions of the UASaro element upstream of a minimal core promoter. Here, we sought to adapt this biosensor to enable ALE.

ARO9 based biosensors have been shown to be induced through exogenous feeding of AAA [2, 48, 147], however they have not been previously used to distinguish differences in endogenously produced AAA. To demonstrate the biosensors capacity to evaluate these differences, we sought to test the biosensor in BY4741, wild-type (WT) yeast and the strain previously developed for muconic acid production containing the *aro3*, *aro4* and *zwf1* genes deleted and the feedback resistant $aro4^{k229l}$ gene integrated (hereafter referred to as ENG). In order to use the biosensor to improve strains for muconic acid production beyond their currently levels, we need to retain all of the beneficial improvements provided by the ENG strain and enable selection capacity beyond the current ENG production level. This requires differentiating WT from ENG

and demonstrating that the biosensor can be further induced through the presence of exogenous AAA.

To test this, we transformed the biosensor plasmid p416-1xUASaro-Leumin-Yecit into the ENG and BY4741 strains and screened with flow cytometry in Complete Synthetic Media (CSM) CSM and CSM with additional AAA. As shown in **Figure 4.1**, the ARO9 based biosensor provides a holistic induction based on the intra and extra-cellular AAA concentration, with the different trends presented by BY4741 and ENG potentially due to differences in basal production. Using the biosensor, the ENG is shown to possess a higher basal expression level and the biosensor can be further induced using exogenous amino acid supplementation demonstrating that it could be used to select for further improvements in endogenous AAA production if used to drive a selectable condition. Having demonstrated that the biosensor could be successfully used to screen improvements in our previously engineered strain, we next turned to converting the biosensor readout from a fluorescent reporter to antibiotic resistance.

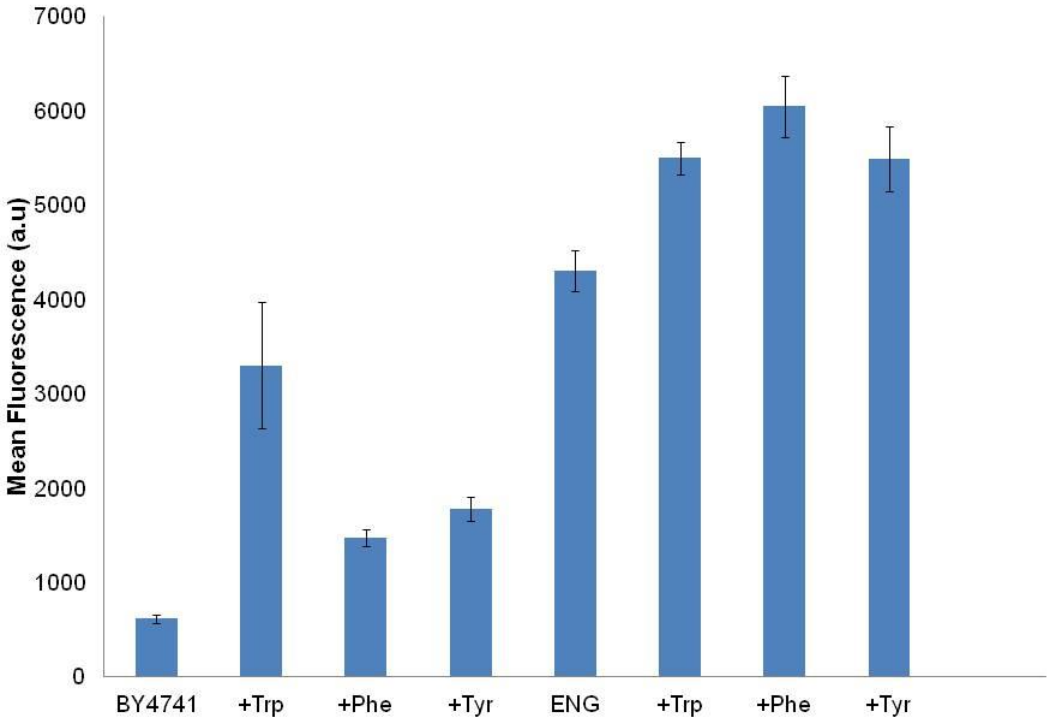

Figure 4.1 Biosensor Inducible Capacity.

The ability of the biosensor to detect differences in endogenous AAA production as well as further inducible capacity is measured. The p416-1xUASaro-Leumin biosensor is tested in BY4741, wild-type (WT) yeast and the strain previously developed for muconic acid production (ENG). All samples were analyzed by flow cytometry on the separate days under the same conditions. Error bars represent standard deviations across biological replicates.

4.3.2 Selective Conditions Analyzed

ALE allows for the selection of mutations which provide a growth benefit in the experimental conditions. Biosensors provide a way of screening beneficial mutations in a high throughput a reporter for screening or selection through resistance. Now that we have demonstrated the ability to detect improvements in AAA production in excess of our previously engineered strain, we used this AAA inducible hybrid promoter to drive expression of the KanNeo gene isolated from the piTY vector [101]. This gene confers weak antibiotic resistance to geneticin (G418) in yeast. The poor resistance conferred by this gene had previously been used to identify tandem-integrations. To create an inducible antibiotic resistance phenotype, we replaced the yECitrine CDS of the p416-4xUASaro-Leumin-Yecit plasmid with the KanNeo G418 resistance gene from the piTY vector to generate the biosensor plasmid p416-4xUASaro-Leumin-KanNeo. This vector was transformed into BY4741 and its growth rate was screened versus a BY4741 expressing a generic yECitrine reporter plasmid to assay the inducible resistance conferred by the KanNeo biosensor as well as the selection potential of anti-metabolites. By employing multiple selective conditions, both biosensor and anti-metabolites, we hope to ensure that the ALE selection pressure results in the selection of improvements in AAA production rather than improvements in biosensor performance.

Using growth rate under selection as a guideline, we tested a number of media conditions to identify the antibiotic and anti-metabolite concentrations which would ensure optimal selection. **Figure 4.2** demonstrates that the 200mg/L G418 slightly

reduces growth rate in both reporter and biosensor strains, but 400mg/L completely abolishes the growth in the control and moderately reduces that of the strain expressing the KanNeo biosensor. Next, we demonstrate that this reduction in growth rate deficit can be alleviated by feeding exogenous AAA to induce the biosensor as compared to our generic reporter plasmid. Finally we tested AAA anti-metabolites previously described in the literature to see if these amino acid analogs could inhibit the growth rate further for use in conjunction with the biosensor. We selected the AAA analogs 4-Fluorophenylalanine (4FP) and 3,4-DL-Dihydroxyphenylalanine (DL-DOPA). These had previously been shown to been toxic and that toxicity could be alleviated through feeding of the natural AAA [160, 161]. The analogs were included this in the media with 400mg/L G418 and the biosensor and control screened for growth, resulting in further growth rate reduction as predicted.

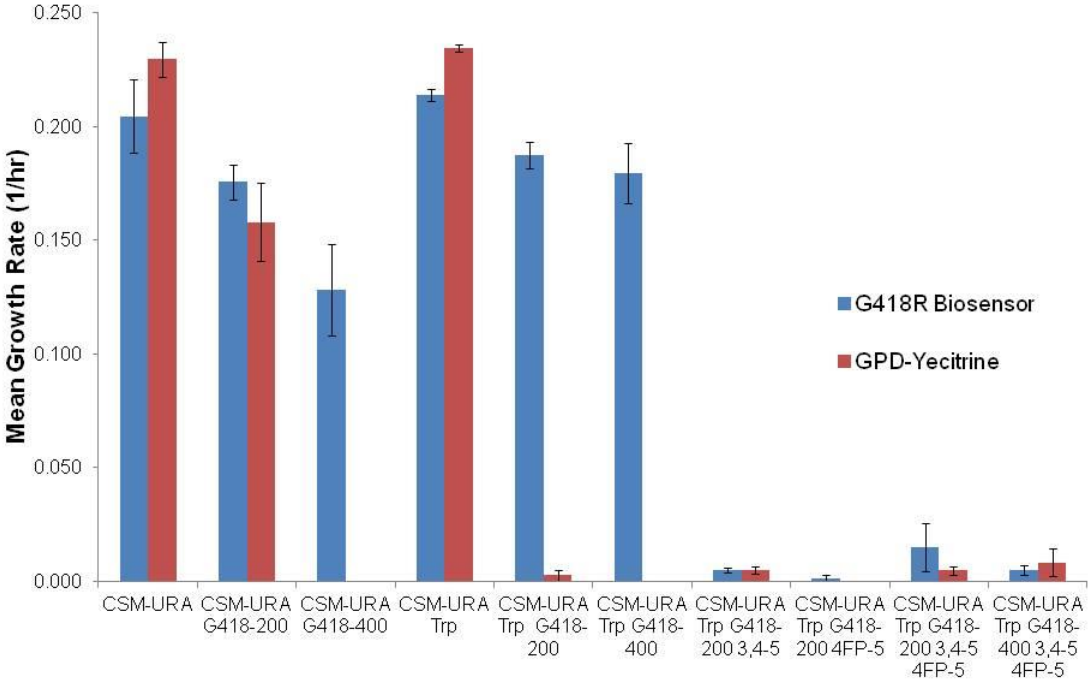

Figure 4.2 Evaluation of media conditions for ALE selection.

Growth rate is measured for WT yeast strains expression 4xUASaro-KanNeo biosensor and control plasmid. The inducible capacity of the 4xUASaro-KanNeo biosensor is demonstrated by recovery of growth rate using exogenous aromatic amino acids as compared to control strain. Further reduction in growth rate is achieved using 3,4 Dihydroxyphenylalanine (5mM) and 4-Fluorophenylalanine (5mM). Error bars represent standard deviations across technical replicates.

4.3.3 Mutation and Long Term Selection for Improved Aromatic Amino Acid Production

While natural mutation rates have been commonly employed in ALE schemes, chemical mutagenesis can greatly speed up the rate of mutation and arrive at desirable evolutionary outcomes quickly with appropriate screening criteria [162]. We used sequential subculturing, transferring a fraction of the population into fresh media to allow improvements in growth rate to be selected for while allowing adaptation to increased concentrations of G418 antibiotic and 4FP.

Using EMS mutagenesis as described previously [163], the ENG strain expressing p416-4xUASaro-KanNeo plasmid was mutagenized resulting in two mutagenenized populations (1x99+% kill and 2x99+% kill) and one unmutagenenized control. This was done in duplicate resulting in six populations which were subcultured at stationary phase into media containing successively greater concentrations of G418 (antibiotic) and 4FP (anti-metabolite). The overall growth trajectory of these six populations (labeled Pop 1-6) is reported in **Figure 4.3**. After 6 rounds of subculturing over 750 hours, the unmutagenized control populations were unable to grow in the selective media conditions, while the mutagenized populations continued to grow in 1000 mg/L G418 and 2mM 4-FP, suggesting that we had reached a point of selection where we believed that a fraction of cells within the successfully growing populations could be producing more AAA than our starting strain.

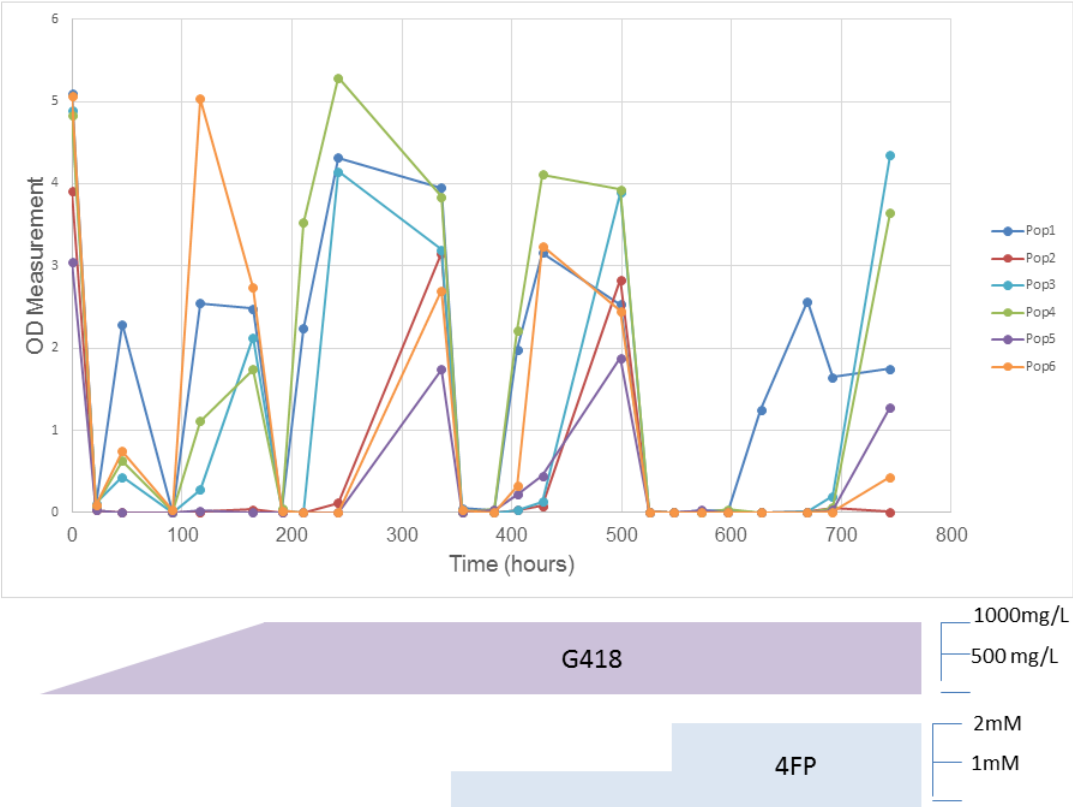

Figure 4.3 Adaptive Laboratory Evolution Log

Following EMS mutagenesis, the six populations (1x and 2x 99% kills and no-EMS control in duplicate) were subcultured in the presence of increasing concentrations of G418 and 4-Fluorophenylalanine. Pop1 and 4 represent 1x 99% kills, while 2 and 4 represent 2x 99% kills and 3 and 6 represent no-EMS respectively. Following 750 hours of subculturing the populations were plated and isolates screened for aromatic amino acid production.

To confirm that our ALE was resulting in the selection of a population of improved AAA producing strains, we plated and isolated colonies from the populations at 750 hours. These colonies were labeled according to the population they were isolated from (i.e. ALE1.01 was the first isolate from Population 1, while ALE 5.02 was the second isolate from Population 5). Colonies were inoculated from plates into CSM and the tyrosine from the supernatant was directly quantified with a plate based tyrosine quantification method utilizing nitrosonapthol- derivatization and compared to the ENG strain as a control [164] presented in **Figure 4.4.** We selected ALE-1.01, ALE-2.08 and ALE-5.01 to generate a second round of diversity through EMS. We cleared the plasmids with 5-FOA to confirm that the plasmids had not integrated. We then retransformed in the p416-4xUASaro-KanNeo vector and repeated the EMS mutagenesis to generate 9 new populations. The 1x and 2x 99% kills or no-EMS controls represented by 1, 2 and 0 in the final digit of the population name. (i.e. ALE1.01.1 was the 1x 99% kill derived from ALE1.01, while Pop-5.01.2 was the 2x 99% kill derived from ALE 5.01). After 575 hours, selection between high performing populations and low was achieved and the AAA production of individual isolates was quantified.

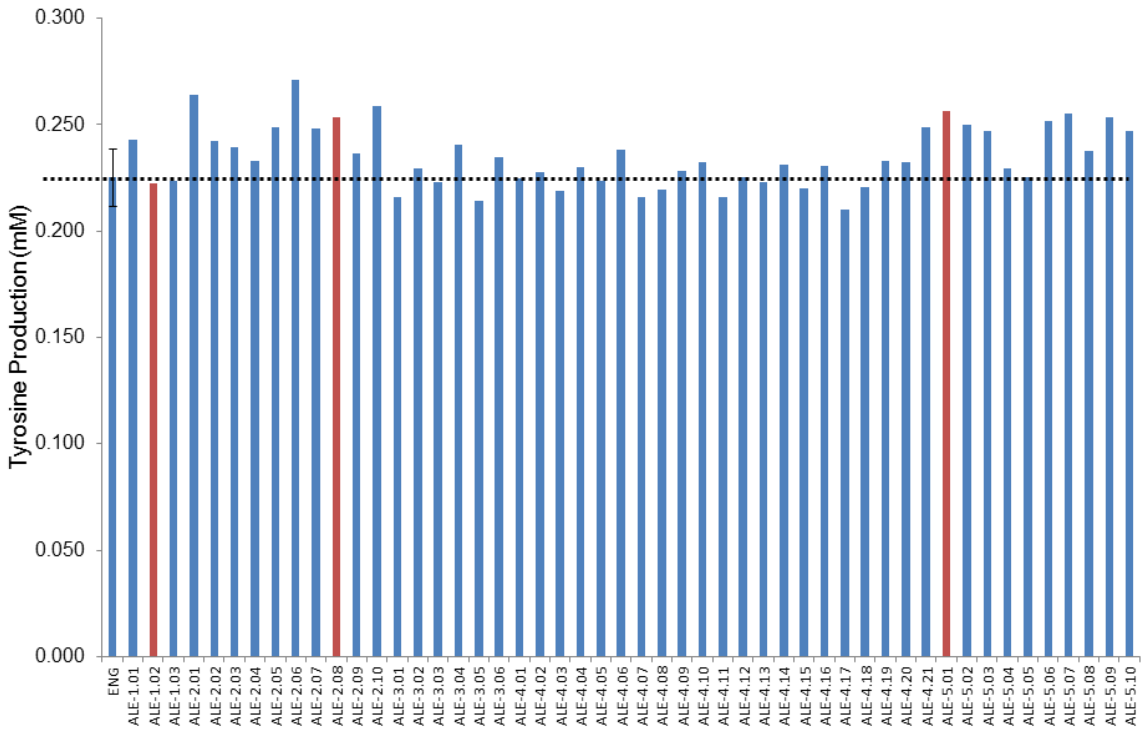

Figure 4.4 Tyrosine Quantification

Tyrosine production of all ALE strains and controls were assayed using the first nitrosonapthol chemical derivatization described in **Chapter 5.4.6** with a high throughput plate reader assay and ENG strain included as a control. All strains were assayed at the same time under the same conditions. The red bars represent the three strains selected for the second round of EMS and ALE. Error bars represent standard deviations across biological replicates. The dotted line represents the mean production of the ENG strain.

We wanted to ensure that this second round would result in improvements, so we started the selection at 1000 mg/L G418 and 1mM 4-FP and over the course of selection, the 4FP concentration was increased to 7.5mM as demonstrated in **Figure 4.5**. This represents a significant improvement in survival over the first round, which failed to

grow after being selected in 1000 mg/L G418 and 2mM 4-FP. After we reached a threshold where differential growth rates were observed between populations, we repeated the process of isolating individual strains from the selected populations. Tyrosine quantification was performed with CSM media containing 2% as well as 4% glucose in the media. The CSM with 4% glucose was selected to test with as it closer mimics the media conditions used to cultivate MuA12 for muconic acid production resulting in the titer of 141mg/L.

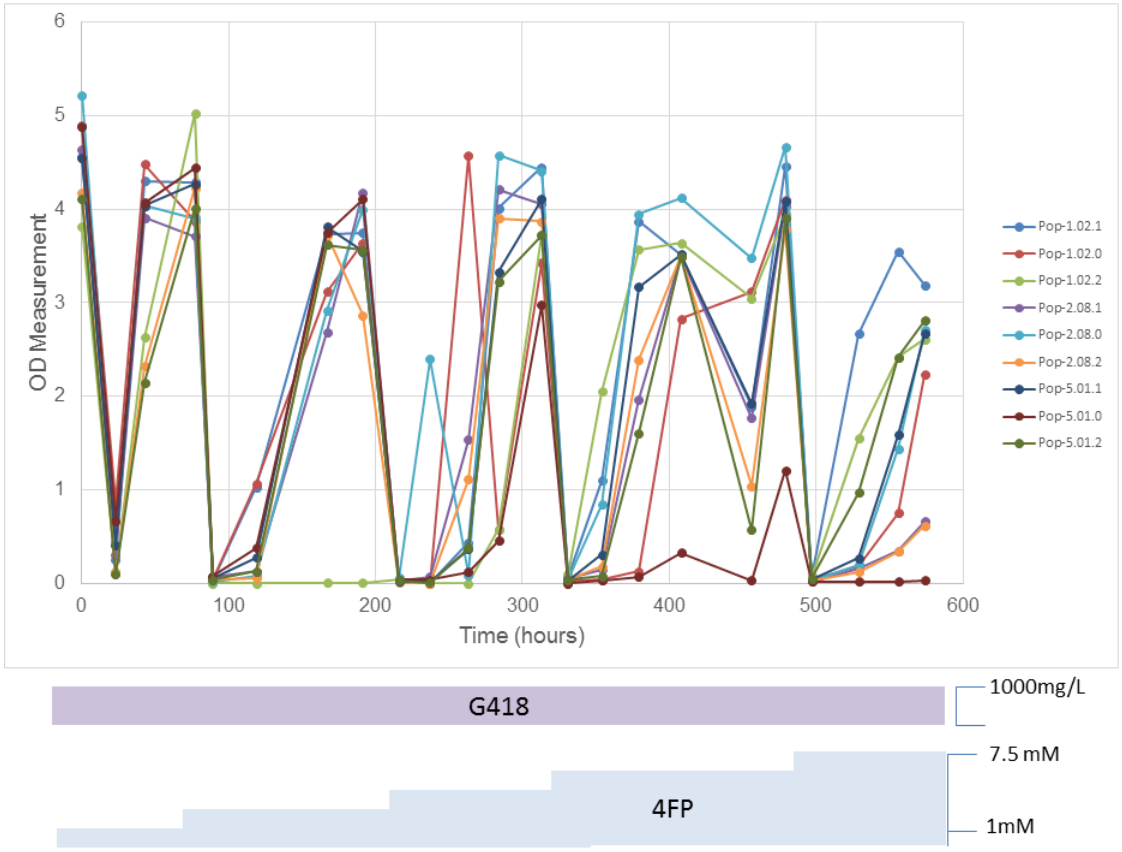

Figure 4.5 Adaptive Laboratory Evolution Log.

Following EMS mutagenesis of the top three strains isolated from the first round of ALE, the 9 populations were subcultured in the presence of 1g/L G418 and increasing concentrations of 4-Fluorophenylalanine. The populations which experienced 1x and 2x 99% kills or no-EMS controls are represented by 1, 2 and 0 in the final digit of the population name. After 575 hours, selection between high performing populations and selection in growth rate was achieved, the AAA production of individual isolates was quantified.

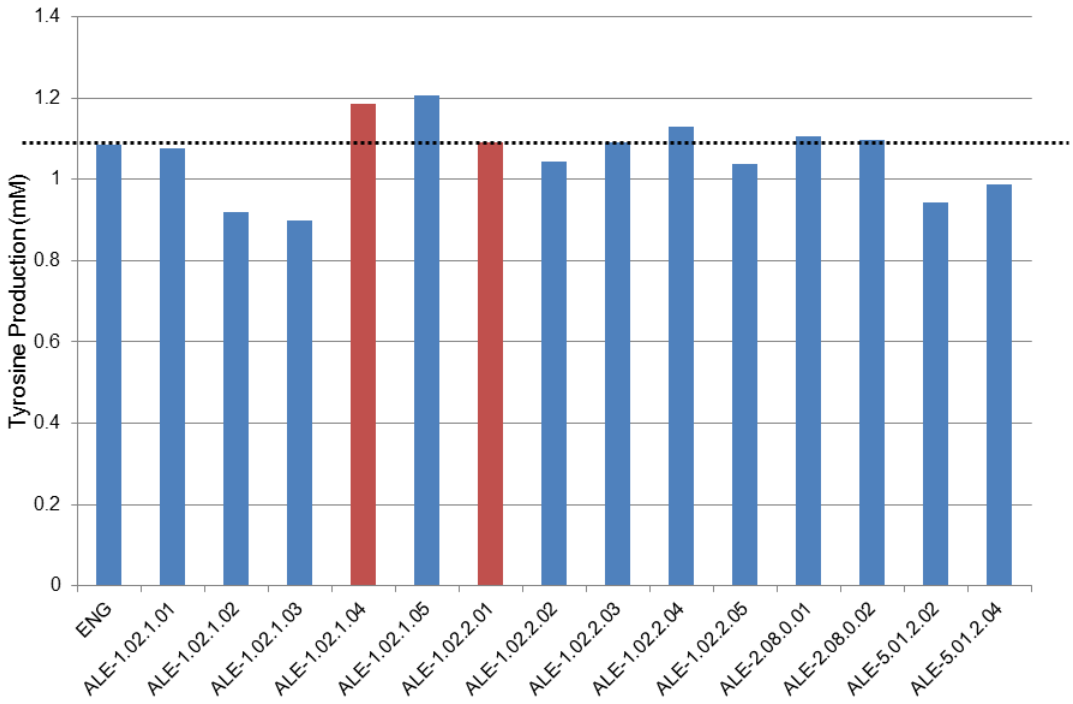

Figure 4.6 Tyrosine Quantification

Tyrosine production of all ALE strains and controls were assayed using the second nitrosonapthol chemical derivatization described in **Chapter 5.4.6** with a high throughput plate reader assay and ENG strain included as a control. All strains were assayed at the same time under the same conditions. The red bars represent the two of the four strains selected for the transformation with the composite pathway. Error bars represent standard deviations across biological replicates. The dotted line represents the production of the ENG strain.

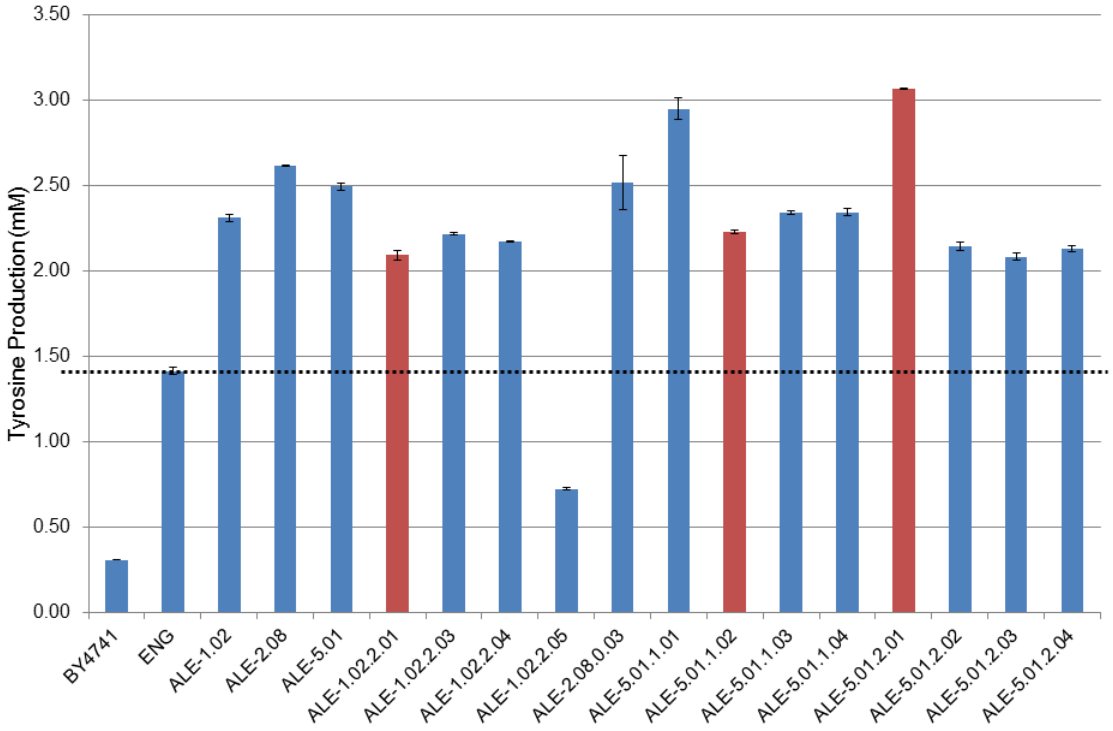

Figure 4.7 Tyrosine Quantification

Tyrosine production of all ALE strains and controls were assayed using the third nitrosonapthol chemical derivatization described in **Chapter 5.4.6** with a high throughput plate reader assay and ENG strain included as a control. All strains were assayed at the same time under the same conditions. The red bars represent the three of the four strains selected for transformation with the composite pathway. Error bars represent standard deviations across biological replicates. The dotted line represents the mean production of the ENG strain.

The tyrosine production described by these experiments clearly demonstrates that we had increased AAA production by 20-100%. However, performance under the different conditions was variable. To improve the accuracy of our measurements we integrated our best practices for tyrosine quantification identified through our three rounds of method development and developed a fourth method. This allowed us to provide a more rigorous, in-depth analysis on what appeared to be some of our best strains by quantifying the tyrosine produced per OD unit in an overnight culture in media lacking amino acids or nitrogen supplementation, reported in **Figure 4.8**. To quickly detect the total AAA concentrations and provide a comprehensive picture of AAA production in these ALE strains, we decided utilize an ARO9 biosensor with a fluorescent reporter as a measurement of total AAA production. We selected the hybrid promoter 5xUASaro-CORE1 containing 5xUASaro elements upstream of the CORE1 minimal synthetic core [2, 149], this was cloned into an EasyClone vector with yECitrine [165] to form p-XII-5xUASaro-CORE1. This vector was integrated into a neutral, high expression locus on Chromosome XII (Jensen 2014) in the seven ALE strains selected for a more in depth analysis were then assayed with flow cytometry. As shown in **Figure 4.9**, the ALE-2.08 and ALE-5.01 did indeed possess increased AAA production despite not seeing an improvement in tyrosine suggesting that mutations were accumulating in regions beneficial to the production of phenylalanine and/or tryptophan.

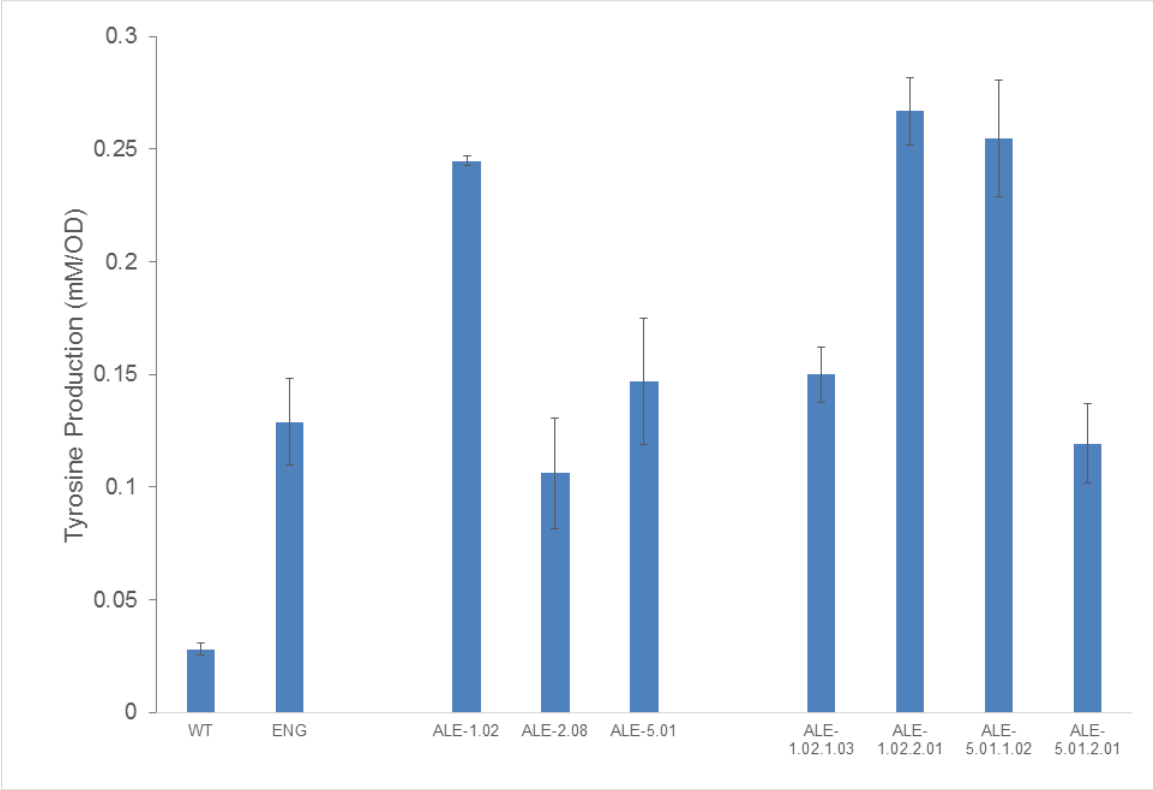

Figure 4.8 Tyrosine production from isolated ALE Strains.

Tyrosine production of all ALE strains and controls were assayed using the fourth nitrosonapthol chemical derivatization described in **Chapter 5.4.6** with a high throughput plate reader assay and ENG strain included as a control. All strains were assayed at the same time under the same conditions. Error bars represent standard deviations across biological replicates.

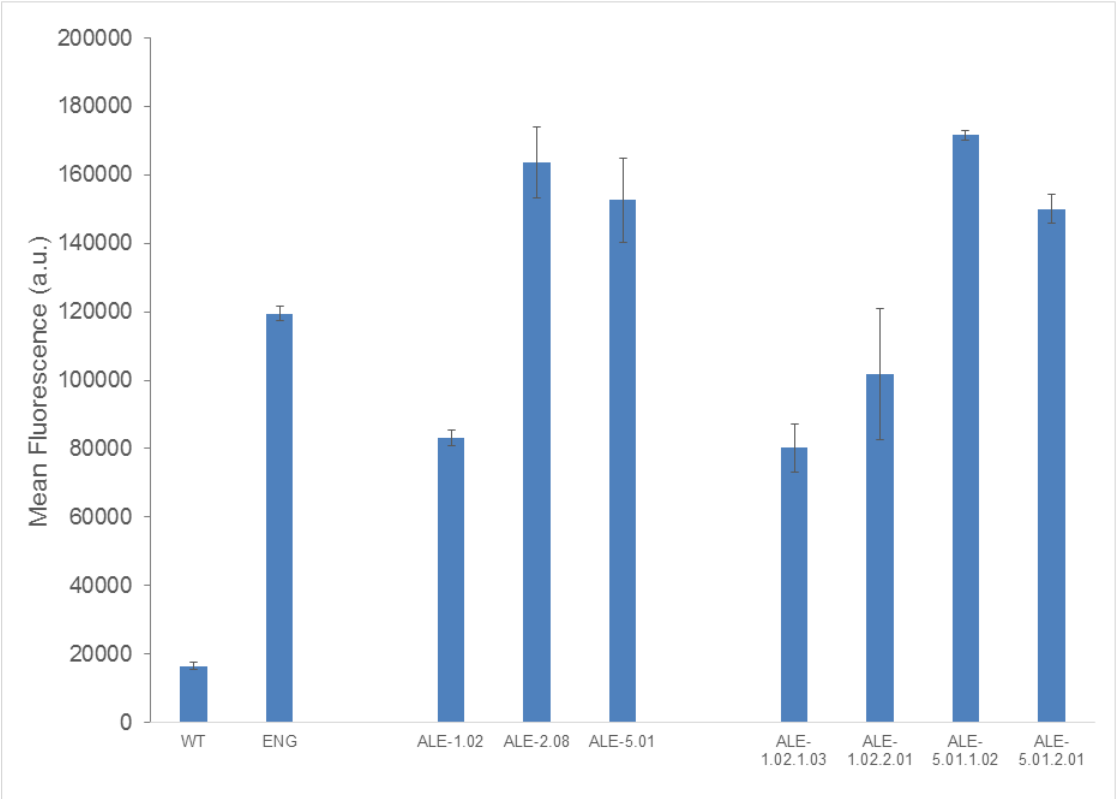

Figure 4.9 Fluorescent based biosensor quantification of Isolated ALE Strains.

The overall AAA production of all ALE strains and controls was assayed will flow cytometry through an integrated biosensor expressing a fluorescent reporter. All strains were assayed at the same time under the same conditions. Error bars represent standard deviations across biological replicates.

Of the four isolated strains, ALE-1.02.1.03 seems to have lost any improvements in AAA production while ALE-1.02.2.01 is enriched in both tyrosine production and AAA as a whole, as measured by the biosensor. The improvement in AAA production presented by these strains confirms that our biosensor based selection platform was successful at isolating improvements in flux through the shikimate pathway. The next

step is confirming that those improvements in flux correlate with improvements in production from our muconic acid composite pathway.

4.3.4 Muconic Acid Production using Evolved Strains

After the final strains have been confirmed for improved AAA flux, we sought to re-divert this flux toward our target muconic acid biosynthetic pathway. To do so, we transformed with the muconic acid pathway into the ALE strains as well as the ENG base strain and their analyzed their production via HPLC. The muconic acid composite pathway draws off of 3-dehydroshikimate (3DHS) and consists of three enzymes previously described: a dehydroshikimate dehydratase from *Podospora anserine* (AROZpodo), protocatechuic acid decarboxylase from *Enterobacter cloacae* (ECL_01944opt), and catechol 1,2-dioxygenase from *Candida albicans* (caHQD2opt) [1].

We first constructed the integration vector PugM, to facilitate a single high-strength integration of ECL_01944opt rather than relying on the inherent variation in tandem integrations used in our previous publication [1] and this high expression ECL_01944opt expression cassette was integrated into the TRP1 locus using the p416-Gal-CAS9 [166]. After confirming integration, the following plasmids were sequentially transformed and confirmed through HPLC analysis: p425-AROZ-HQD2, p426-AROZ-ECL_01944opt and p413-TKL1 to create strains MuA-1.02.1.04, MuA-1.02.2.01, MuA-5.01.1.02 and MuA-5.01.2.01. The production from this composite pathway in the ALE

strains and the previously engineered ENG strain (MuA13) is seen in **Figure 4.10**. This resulted in 3 fold composite pathway production relative to the ENG strain. This demonstrates the high AAA flux due to the ALE can be translated to improvements in output from our composite pathway. We next want to re-route flux away from the downstream products, AAA, and into the composite pathway and on to improving muconic acid titers through local pathway re-wiring.

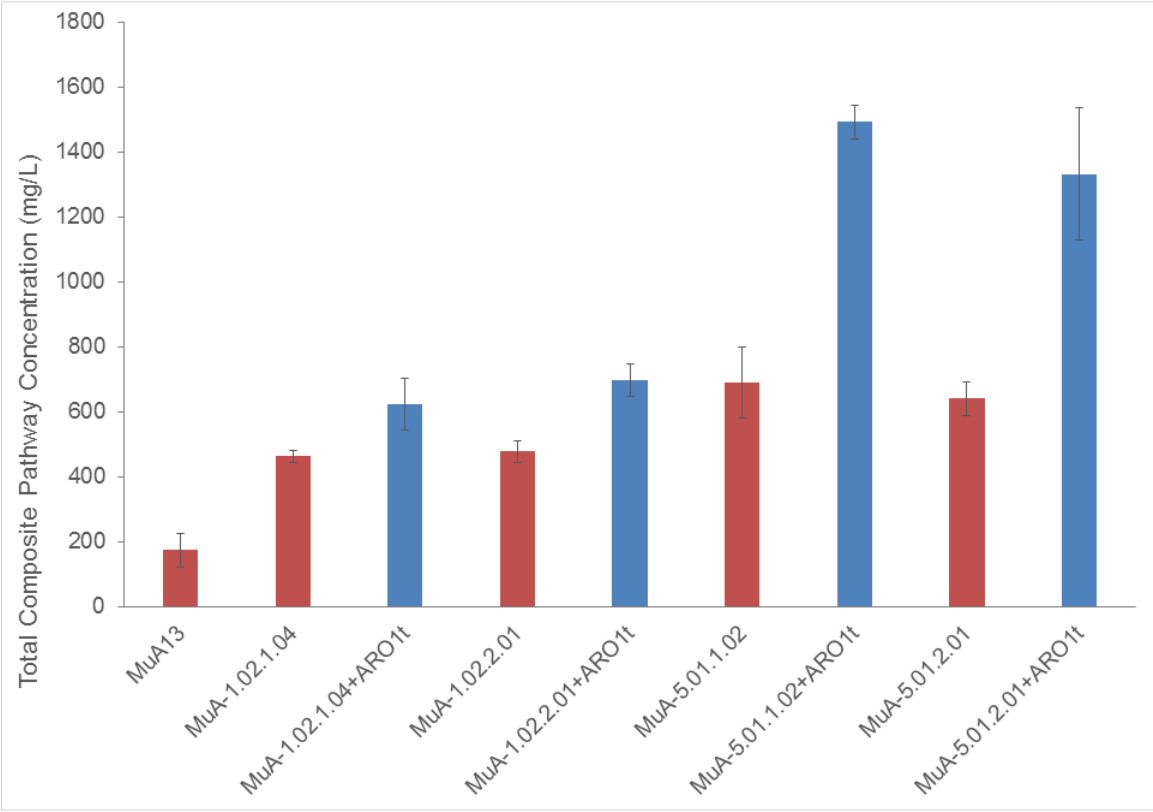

Figure 4.10 Composite Pathway Production of ALE Strains.

The red bars represent ALE and control strains expressing the muconic acid composite pathway, while MuA13 represents ENG with similar pathway. Strains were cultured in the flask and total muconic acid pathway production was quantified by HPLC. Blue bars represent ALE strains with overexpression of the truncated ARO1t protein rerouting flux into the composite pathway. Error bars represent standard deviations across biological replicates.

4.3.5 ARO1 Truncation

To reroute carbon flux and divert this flux for producing muconic acid in the ALE strains, the flux needed to be rerouted to DHS and on to the composite pathway. The

Shikimate pathway in yeast is largely consolidated into one enzyme, the pentafuncational ARO1p. Other groups have addressed this either by cutting out ARO1p entirely and replacing it with the orthoganol pathway from *E. coli* [17], or through point mutations based on homology to *E. coli* proteins [24]. Here we propose an alternative strategy to engineering ARO1p to increase flux into 3DHS while limiting production of shikimic acid. Previous homology modeling has identified the 1306-1588 residues of ARO1p to correspond with the shikimate dehydrogenase from *E. coli*, *aroE*. We proposed that by eliminating the enzymatic function of *ARO1d* through a truncation we would be able to direct flux directly into 3DHS at the expense of the downstream AAA. [99, 114]

To identify a functional ARO1p truncation (ARO1t), we cloned mutant *ARO1*genes on plasmids and expressed them in *aro1Δ* alongside plasmids harboring the *aroE* gene from *E. coli* and screened for growth in media lacking AAA supplementation. While the simple truncation at codon 1305 failed to facilitate growth, the addition of 40 amino acids of *ARO1d* facilitated growth in complementation with *aroE*. This suggests that these additional residues might be required to facilitate folding or other tasks essential to proper enzyme function. The functional ARO1t gene was then cloned into the p424 vector and transformed into ALE strains expressing the composite pathway. The total pathway potential from these strains was then assayed and titers are reported in **Figure 4.10**, with MuA-5.01.1.02+ARO1t capable of producing roughly 1.5 g/L which is 7.5 fold of the MuA13 and a significant improvement through re-routing flux with the ARO1t.

4.3.6 Composite Pathway Optimization

With ARO1t introduced into the MuA-5.01.1.02 strain, MuA-5.01.1.02+ARO1t was able to produce over 1.5g/L from the composite pathway at the flask scale; however, this failed to correlate with high concentrations of muconic acid due to the poor enzymatic activity of AroY. Since our original work in this area, other groups have discussed the poor enzymatic activity of AroY and it remains a rate limiting step in muconic acid production [17, 24]. While Suastegui and coworkers demonstrated that this can be somewhat alleviated through control of oxygenation, we propose improving conversion through the use of an alternative PCA decarboxylase. Recent work identified scPAD1 as a functional decarboxylase in the *S. cerevisiae* genome, natively used to detoxify cinnamic acid [167]. To improve the conversion of PCA into muconic acid, we overexpressed the endogenous scPAD1 gene in the MuA-5.01.1.02+ARO1t strain, resulting in significant improvements in throughput into the final muconic acid product with over 550mg/L muconic acid produced in shake flask and comparable to the titers Suastegui and coworkers report at the bioreactor scale. The resulting muconic acid titer is compared to the previously reported MuA12 strain in **Figure 4.11** demonstrating the resulting four fold improvement in titer and yield provided by ALE and pathway optimization.

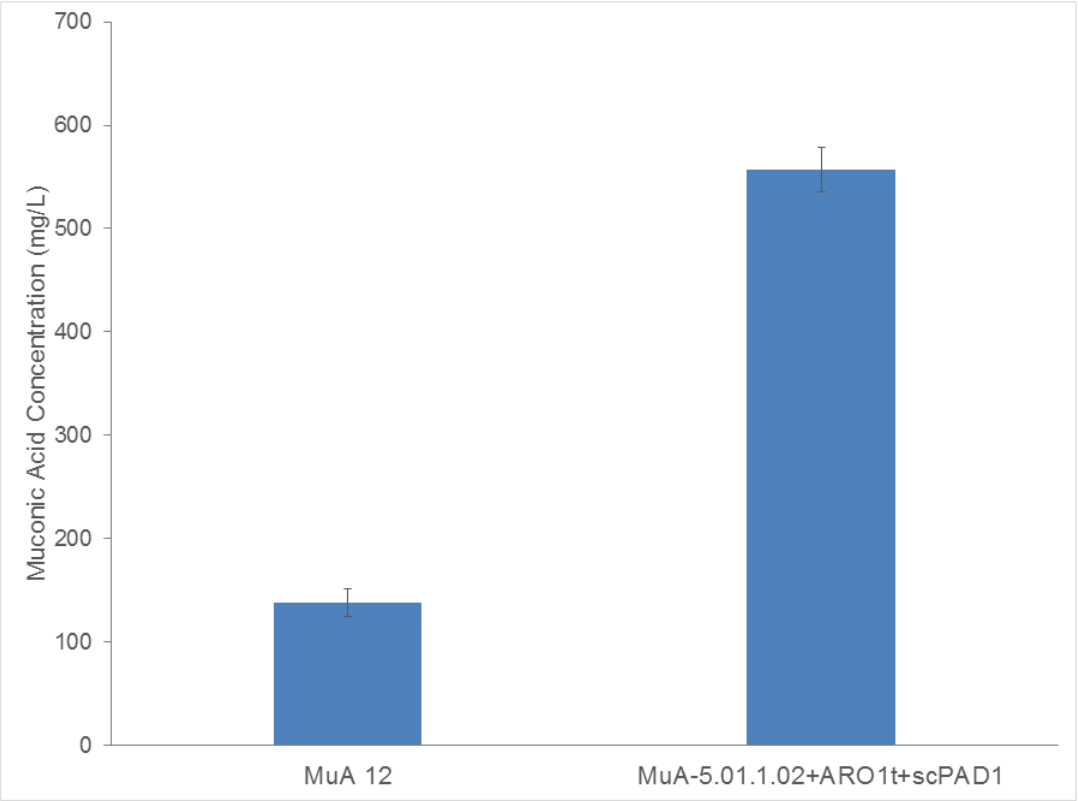

Figure 4.11 Muconic Acid Production of MuA-5.01.1.02+ARO1t+scPAD1 Strain.

Muconic acid production of the MuA-5.01.1.02+ARO1t+scPAD1 strain in flask compared against our previously reported muconic acid producing strain MuA12. MuA-5.01.1.02+ARO1t+scPAD1 integrates our initial metabolic engineering work, ALE and final pathway rewiring through ARO1t and scPAD1 overexpression demonstrating a 4-fold improvement in yield and titer. Error bars represent standard deviations across biological replicates.

4.3.7 Bioreactor Fermentation

While this strain produced the highest muconic acid to date in *S. cerevisiae* at flask scale, we desired to scale up the strain and demonstrate further improvements through control of pH, oxygenation and media formulation which have previously been shown to have great impact on final titer and yield of aromatic compounds [24, 152]. We selected CSM for our media formulation to closer replicate the media conditions which the ALE strains were evolved in. To closely mimic the microaerobic conditions previously shown to improve PCA decarboxylase activity, we maintained an SLPM of 0.5 and maintained pH control at 5.0. The bioreactor fermentation was operated as a fed-batch process with 3 media feedings throughout the run when the majority of glucose had been consumed. As shown in **Figure 4.12**, this process resulted in the production of 1.94 g/L of muconic acid after 11 days representing the highest titer ever reported from glucose of muconic acid in *S. cerevisiae* and of any product from the shikimate pathway.

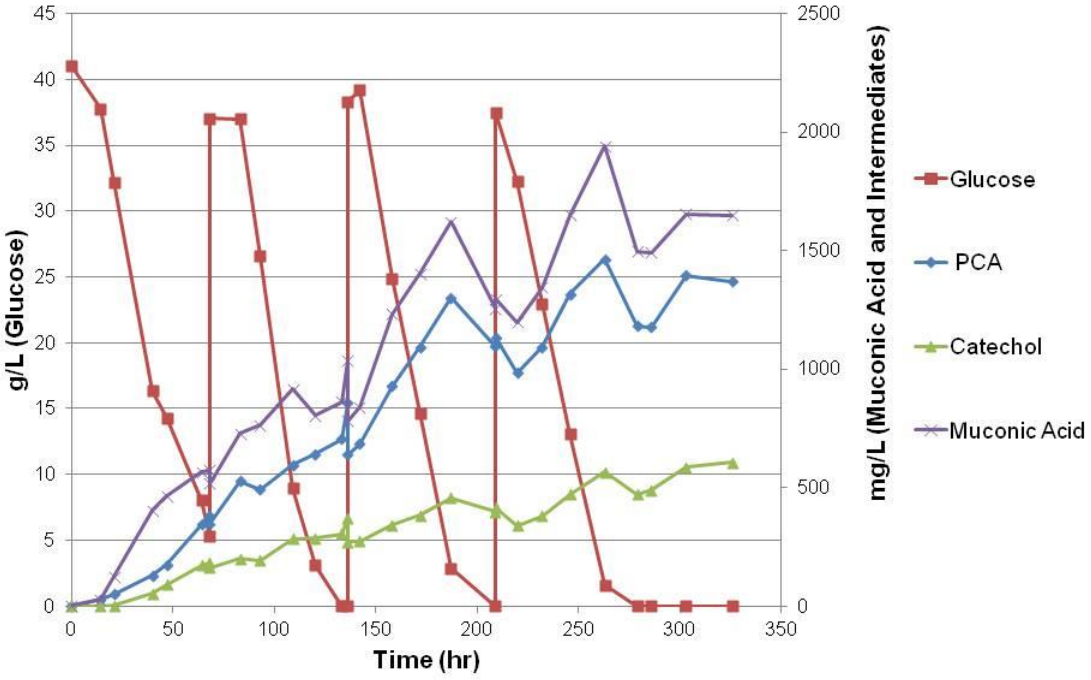

Figure 4.12 Bioreactor Fermentation.

Final muconic acid producing strain MuA-5.01.1.02+ARO1t+scPAD1 was cultured in bioreactor under batch conditions and PCA, Catechol, total muconic acid and glucose quantified. Three times throughout the run, as glucose was depleted additional media was spiked in, facilitating additional cell growth and production.

4.4 Concluding Remarks

This chapter represents a generic strategy which could be employed other biosensors to select for improved production of other metabolites. After our initial success expressing the muconic acid composite pathway and engineering the strain

through rational metabolic engineering, we then developed and employed an evolutionary strategy utilizing a biosensor to direct flux through the shikimate pathway employing AAA as a surrogate for muconic acid. Following two iterations of mutation and selection, we isolated strains of yeast capable of double the AAA production and three fold of the muconic acid pathway potential. We then performed some local pathway rewiring to ensure that our improvements flux through the shikimate pathway could be successfully re-routed into the composite pathway.

To reduce flux downstream of 3DHS, we first engineered a truncated ARO1p which removed the shikimate dehydrogenase activity from the *ARO1d* domain. We then overexpressed this protein in our ALE strains expressing the muconic acid composite pathway resulting in strains capable of 7.5 fold output from the composite pathway and roughly doubling titers in our best performing strains. The final step in strain engineering was improving conversion from PCA to muconic acid through the expression of the endogenous PCA decarboxylase, scPAD1, which resulted in a strain capable of over 550mg/L muconic acid production in shake flask and 1.94 g/L in fed-batch bioreactor with pH and DO control. This represents a nearly a 14-fold improvement over our previously reported strain, which is 4-fold of the highest reported titer of muconic acid in yeast. Through ALE and local pathway optimization we were able to accomplish the highest production of muconic acid in yeast, as well as one of the highest reporter titer of a shikimate pathway derivative.

Plasmid Name	Source of Assembly Method	Primers
p416-1xUASaro-Leumin-yECitine	[2]	n/a
p416-4xUASaro-Leumin-yECitine	[2]	n/a
p416-4xUASaro-Leumin-KanNeo	Restriction Cloning with XbaI and SalI	Fwd: CTAGTCTAGAATGAGCCATATTCAACGG
		Rev: ACGCGTCGACGAAAAACTCATCGAGCATCA
p416-GPD-yECitrine	[62]	n/a
p-XII-5xUASaro-CORE1-yECitrine	Restriction Cloning with BamHI and PmeI	Fwd: Ggggtaccgatcgcgtcagctgaagctt
		Rev: CGGGATCCatcgcacgcattccgttg
pug6	[123]	n/a
pugM-UASCLB–UASCIT–UASTEF-GPD-ECL_01944opt-Tprm9	Gibson Assembly	Vector Fwd: CAACGCTTCGGAAAATACGATGTTGAAAATccgcg gatctgccggtctccctatagtgag
		Vector Rev: aaaaaaggagtagaaacattttgaagctatgatatcacctaataacttcgtatagca tac
		Promoter Fwd: gtatgctatacgaagttattaggtgatatc-atagcttcaaaatgtttctactccttttt
		Promoter Rev: ATTTATTGGGTTTTGCATactagttctaga-atccgtcgaaactaagttctggtgttttaa
		ECL_01944opt Fwd: ttaaaacaccagaacttagtttcgacggat-tctagaactagtATGCAAAACCCAATAAAT
		ECL_01944opt Rev: GCTAGTGTCTCCCGTCTTCTGT-GGCGCGCC-CTATTTTTTGTCAGAAAATAATTCAGGGGC
		Tprm9 Fwd: GCCCCTGAATTATTTTCTGACAAAAAATAG-GGCGCGCC-ACAGAAGACGGGAGACACTAGC
		Tprm9 Rev: ctcactatagggagaccggcagatccgcggATTTTCAACATCGTA TTTTCCGAAGCGTTG
		Trp1 Integration Fwd: AATTTCACAGGTAGTTCTGGTCCATTGGTGAAAG TTTGCGGCTTGCAGAGCACAGAGGCCGCAGAAT GTtgcaggtcgacaacccttaat
		Trp1 Integration Rev: AATTTGCTATTTTGTTAGAGTCTTTTACACCATTT GTCTCCACACCTCCGCTTACATCAACACCAATTT TCAACATCGTATTTTCCGAAG
p416-Gal-Cas9-Trp1	[166]	Trp1 gRNA Sequence: AGGAACTCTTGGTATTCTTG

Table 4.1: continued next page.

p425-TEF-pa5_5120opt/GPD-caHQD2opt	[1]	n/a
p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9	Gibson Assembly	Vector Rev: aaaaaaggagtagaaacattttgaagctatTCCCTTTAGTGAGGGTTAATTGCG
		Vector Fwd: TTCGGAAAATACGATGTTGAAAATggtaccCCCTATAGTGAGTCGTATTACGCG
		Vector Fwd: TTCGGAAAATACGATGTTGAAAATggtaccCCCTATAGTGAGTCGTATTACGCG
		pa Cassette Rev: tgaaatggcgagtattgataatgataaactGAGCTCCAAATTAAAGCCTTCG
		Ecl Cassette Fwd: gggacgctcgaaggctttaatttggagctcAGTTTATCATTATCAATACTCGCCATTTC
		Ecl Cassette Rev: cgcgtaatacgactcactataggggtaccATTTTCAACATCGTATTTTCCGAAGCG
p413-TEF-scTKL1	[1]	n/a
p415-TEF-ecAroE	Restriction Cloning with SpeI and XhoI	Fwd: GgactagtATGGAAACCTATGCTGTTTTTGGTAATC
		Rev: CCGctcgagTCACGCGGACAATTCCTCC
p413-TEF-ARO1t	Restriction Cloning with SpeI and XhoI	Fwd: GGactagtATGGTGCAGTTAGCCAAAGTCC
		Rev: CCGctcgagCTATAAAATTTCATAGCCAGTGTTATGTAAAATTGG
p424-GPD-ARO1t	Restriction Cloning with SpeI and XhoI	Fwd: GGactagtATGGTGCAGTTAGCCAAAGTCC
		Rev: CCGctcgagCTATAAAATTTCATAGCCAGTGTTATGTAAAATTGG
scPAD1-Entry	BsmbI Golden Gate	Fwd: gcatcgtctcatcggtctcatATGCTCCTATTTCCAAGAAGAACTAATATAGC
		Rev: atgccgtctcaggtctcaggatTTACTTGCTTTTTATTCCTTCCCAACGAG
scPAD1-IV	BsaI Golden Gate	Part vectors used: scPAD1 Entry vector, pYTK002,9,53,72,78,87,90,93

Table 4.1: Plasmids used in this chapter.

Plasmids were derived from those described in [122] with the exception of p-XII-5xUASaro-CORE1 and scPAD1-IV.

Strain Name	Plasmids	Genotype Features	Source
BY4741		*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0*	Euroscarf Y00000
BY4741-p416-1xUASaro-yecitrine		*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0*	(Leavitt, J. M., et al., 2016)
ENG		*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ*	MuA10 w/o AROY
ENG-p416-1xUASaro-yecitrine	p416-1xUASaro-yecitrine	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ*	Plasmid Transformation
BY4741-p416-4xUASaro-KanNeo	p416-4xUASaro-KanNeo		Plasmid Transformation
BY4741-p416-GPD-Yecitrine	p416-GPD-Yecitrine		Plasmid Transformation
ENG-p416-4xUASaro-KanNeo	p416-4xUASaro-Leumin-KanNeo	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ*	Plasmid Transformation
Pop 1	p416-4xUASaro-Leumin-KanNeo		EMS 1x99% Kill
Pop 2	p416-4xUASaro-Leumin-KanNeo		EMS 2x99% Kill
Pop 3	p416-4xUASaro-Leumin-KanNeo		Control1
Pop 4	p416-4xUASaro-Leumin-KanNeo		EMS 1x99% Kill
Pop 5	p416-4xUASaro-Leumin-KanNeo		EMS 2x99% Kill
Pop 6	p416-4xUASaro-Leumin-KanNeo		Control2
ALE-1.01	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02	p416-4xUASaro-Leumin-KanNeo		
ALE-1.03	p416-4xUASaro-Leumin-KanNeo		
ALE-2.01	p416-4xUASaro-Leumin-KanNeo		
ALE-2.02	p416-4xUASaro-Leumin-KanNeo		
ALE-2.03	p416-4xUASaro-Leumin-KanNeo		
ALE-2.04	p416-4xUASaro-Leumin-KanNeo		
ALE-2.05	p416-4xUASaro-Leumin-KanNeo		
ALE-2.06	p416-4xUASaro-Leumin-KanNeo		

Table 4.2: continued next page.

ALE-2.07	p416-4xUASaro-Leumin-KanNeo		
ALE-2.08	p416-4xUASaro-Leumin-KanNeo		
ALE-2.09	p416-4xUASaro-Leumin-KanNeo		
ALE-2.10	p416-4xUASaro-Leumin-KanNeo		
ALE-3.01	p416-4xUASaro-Leumin-KanNeo		
ALE-3.02	p416-4xUASaro-Leumin-KanNeo		
ALE-3.03	p416-4xUASaro-Leumin-KanNeo		
ALE-3.04	p416-4xUASaro-Leumin-KanNeo		
ALE-3.05	p416-4xUASaro-Leumin-KanNeo		
ALE-3.06	p416-4xUASaro-Leumin-KanNeo		
ALE-4.01	p416-4xUASaro-Leumin-KanNeo		
ALE-4.02	p416-4xUASaro-Leumin-KanNeo		
ALE-4.03	p416-4xUASaro-Leumin-KanNeo		
ALE-4.04	p416-4xUASaro-Leumin-KanNeo		
ALE-4.05	p416-4xUASaro-Leumin-KanNeo		
ALE-4.06	p416-4xUASaro-Leumin-KanNeo		
ALE-4.07	p416-4xUASaro-Leumin-KanNeo		
ALE-4.08	p416-4xUASaro-Leumin-KanNeo		
ALE-4.09	p416-4xUASaro-Leumin-KanNeo		
ALE-4.10	p416-4xUASaro-Leumin-KanNeo		
ALE-4.11	p416-4xUASaro-Leumin-KanNeo		
ALE-4.12	p416-4xUASaro-Leumin-KanNeo		
ALE-4.13	p416-4xUASaro-Leumin-KanNeo		
ALE-4.14	p416-4xUASaro-Leumin-KanNeo		
ALE-4.15	p416-4xUASaro-Leumin-KanNeo		
ALE-4.16	p416-4xUASaro-Leumin-KanNeo		
ALE-4.17	p416-4xUASaro-Leumin-KanNeo		
ALE-4.18	p416-4xUASaro-Leumin-KanNeo		
ALE-4.19	p416-4xUASaro-Leumin-KanNeo		
ALE-4.20	p416-4xUASaro-Leumin-KanNeo		
ALE-4.21	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01	p416-4xUASaro-Leumin-KanNeo		
ALE-5.02	p416-4xUASaro-Leumin-KanNeo		
ALE-5.03	p416-4xUASaro-Leumin-KanNeo		
ALE-5.04	p416-4xUASaro-Leumin-KanNeo		
ALE-5.05	p416-4xUASaro-Leumin-KanNeo		
ALE-5.06	p416-4xUASaro-Leumin-KanNeo		
ALE-5.07	p416-4xUASaro-Leumin-KanNeo		
ALE-5.08	p416-4xUASaro-Leumin-KanNeo		
ALE-5.09	p416-4xUASaro-Leumin-KanNeo		
ALE-5.10	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02			5-FoA

Table 4.2: continued next page.

ALE-2.08			5-FoA
ALE-5.01			5-FoA
Pop-1.02.1	p416-4xUASaro-Leumin-KanNeo		EMS 1x99% Kill
Pop-1.02.2	p416-4xUASaro-Leumin-KanNeo		EMS 2x99% Kill
Pop-1.02.0	p416-4xUASaro-Leumin-KanNeo		Control
Pop-2.08.1	p416-4xUASaro-Leumin-KanNeo		EMS 1x99% Kill
Pop-2.08.2	p416-4xUASaro-Leumin-KanNeo		EMS 2x99% Kill
Pop-2.08.0	p416-4xUASaro-Leumin-KanNeo		Control
Pop-5.01.1	p416-4xUASaro-Leumin-KanNeo		EMS 1x99% Kill
Pop-5.01.2	p416-4xUASaro-Leumin-KanNeo		EMS 2x99% Kill
Pop-5.01.0	p416-4xUASaro-Leumin-KanNeo		Control
ALE-1.02.1.01	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.1.02	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.1.03	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.1.04	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.1.05	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.2.01	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.2.02	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.2.03	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.2.04	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.2.05	p416-4xUASaro-Leumin-KanNeo		
ALE-2.08.0.01	p416-4xUASaro-Leumin-KanNeo		
ALE-2.08.0.02	p416-4xUASaro-Leumin-KanNeo		
ALE-2.08.0.03	p416-4xUASaro-Leumin-KanNeo		
ALE-2.08.0.04	p416-4xUASaro-Leumin-KanNeo		
ALE-2.08.0.05	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.1.01	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.1.02	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.1.03	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.1.04	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.2.01	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.2.02	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.2.03	p416-4xUASaro-Leumin-KanNeo		
ALE-5.01.2.04	p416-4xUASaro-Leumin-KanNeo		
ALE-1.02.1.04			5-FoA
ALE-1.02.2.01			5-FoA
ALE-5.01.1.02			5-FoA
ALE-5.01.2.01			5-FoA
ALE-1.02		*XII::5xUASaro-CORE1-yECitrine*	Integration
ALE-2.08		*XII::5xUASaro-CORE1-yECitrine*	Integration
ALE-5.01		*XII::5xUASaro-CORE1-yECitrine*	Integration

Table 4.2: continued next page.

ALE-1.02.1.04		*XII::5xUASaro-CORE1-yECitrine*	Integration
ALE-1.02.2.01		*XII::5xUASaro-CORE1-yECitrine*	Integration
ALE-5.01.1.02		*XII::5xUASaro-CORE1-yECitrine*	Integration
ALE-5.01.2.01		*XII::5xUASaro-CORE1-yECitrine*	Integration
MuA13	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	ENG
MuA-1.02.1.04	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	ALE-1.02.1.04
MuA-1.02.2.01	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	ALE-1.02.2.01
MuA-5.01.1.02	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	ALE-5.01.1.02
MuA-5.01.2.01	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	ALE-5.01.2.01
MuA-1.02.1.04+ARO1t	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1, p424-ARO1t	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	MuA-1.02.1.04
MuA-1.02.2.01+ARO1t	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1, p424-ARO1t	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	MuA-1.02.2.01

Table 4.2: continued next page.

MuA-5.01.1.02+ARO1t	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1, p424-ARO1t	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	MuA-5.01.1.02
MuA-5.01.2.01+ARO1t	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1, p424-ARO1t	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9*	MuA-5.01.2.01
MuA-5.01.1.02+ARO1t+scPAD1	p425-TEF-pa5_5120opt/GPD-caHQD2opt, p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt -Tprm9, p413-TEF-scTKL1, p424-ARO1t	*Mat a; his Δ1; leu2Δ0; met15Δ0; ura3Δ0; aro3Δ; aro4Δ::PGPD-aro4k229l; zwf1Δ; trp1Δ::UASCLB–UASCIT–UASTEF-GPD-ECL_01944optTprm9 ; leu2:: scPAD1*	MuA-5.01.1.02+ARO1t

Table 4.2: Yeast strains used in this chapter.

A table of yeast used in this chapter, their lineage, plasmids harbored and known genotypic characteristics.

Chapter 5: Materials and Methods

5.1 Common Materials and Methods

5.1.1 Strains and media

Saccharomyces cerevisiae strain BY4741 (*Mat a; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*) was used as the primary host strain for this work (obtained from EUROSCARF). Yeast strains were routinely propagated at 30°C in Yeast Extract Peptone Dextrose (YPD) medium, yeast synthetic complete (YSC) medium, or yeast synthetic minimal (YSM) medium. YPD medium is composed of 10 g/L yeast extract, 20 g/L peptone, and 20 g/L glucose. YSC medium is composed of 6.7 g/L yeast nitrogen base, 20 g/L glucose, and either CSM-Ura, CSM-His, CSM-Leu, CSM-Trp or combination thereof (MP Biomedicals, Solon, OH), depending on the required auxotrophic selection. YSM medium is composed of 6.7 g/L yeast nitrogen base, 20 g/L glucose, 20 mg/L methionine, and 10 mg/L adenine. *Escherichia coli* strain *DH10β* was used for all cloning and plasmid propagation. *DH10β* was grown at 37 °C in Luria-Bertani (LB) broth supplemented with 50 μg/mL of ampicillin. All strains were cultivated with 225 RPM orbital shaking. Yeast and bacterial strains were stored at -80°C in 15% glycerol

5.1.2 Plasmid construction

Standard cloning and bacterial transformations were performed according to Sambrook and Russell [168]. Genomic DNA from *S. cerevisiae* and *E. coli* were obtained using Wizard Genomic DNA Extraction Kit from Promega. PCR reactions used Phusion High-Fidelity DNA Polymerase from New England Biolabs (Ipswich, MA) and followed supplier instructions; primers were purchased from Integrated DNA

Technologies (Coralville, Iowa). Antarctic phosphatase and all restriction enzymes were purchased from New England Biolabs (Ipswich, MA). Fermentas T4 DNA ligase and all other enzymes and chemicals were purchased through Thermo Fisher Scientific (Waltham, MA). Vectors were isolated using the Zyppy Plasmid Miniprep kit from Zymo Research Corp. (Irvine, CA) and DNA purification was performed with a Qiaquick PCR Cleanup kit (Qiagen, Valencia, CA). Some cloning procedures required gel extraction, which was accomplished with the Fermentas GeneJET Gel Extraction Kit from Thermo Fisher Scientific (Waltham, MA). All plasmids and genes were sequenced confirmed to ensure correct identify of the insert prior to yeast transformations

5.2 Materials and Methods for Chapter 2

5.2.1 Plasmid construction

With the exception of integration vectors, all plasmids were constructed using standard yeast plasmids with either the *TEF2* or *GPD* (*TDH3*) promoter from {Mumberg, 1995 #68} (**Table 2.1**, at the end of the chapter). The following genes: *AroZ* and *AroY* from *Klebsiella pneumoniae*, *CatA* from *Acinetobacter baylyi*, *QutC* from *Aspergillus niger*, Pa_5_5120, Pa_0_880, and Pa_4_4540 from *Podospora anserina*, ECL_01944 from *Enterobacter cloacae*, and *HQD2* from *Candida albicans* were codon-optimized for expression in *S. cerevisiae* and synthesized by Blue Heron Biotechnology (Bothell, WA). These genes were either gel extracted or cloned via PCR (see **Table 2.1** for primers) and inserted into the desired plasmid mulitcloning site using the XbaI or SpeI site at the 5'

end of the gene and the ClaI, EcoRI or SalI site at the 3' end of the gene. The *FDC1*, *PAD1* and *ARO4* genes from *S. cerevisiae* and DEHA2F15906g, DEHA2G00682g and DEHA2C14806g from *Debaryomyces hansenii* were cloned via PCR from extracted gDNA (obtained using the Wizard Genomic DNA Extraction Kit from Promega, Madison, WI). The wildtype *K. pneumoniae AroZ* and *AroY* genes and *A. baylyi CatA* gene were cloned from *E. coli* expression plasmids provided by Draths Corporation.

When it was desired to express two genes on a single plasmid, each was first cloned into a separate plasmid as described above, and then the expression cassette containing one of the genes, including the surrounding promoter and terminator, was cloned into a single restriction site on the other plasmid.

The feedback-resistant *ARO4* mutant, $aro4_{k229l}$, was created by cloning the wildtype gene into a vector to create p416-TEF-sc*ARO4* and then using the QuikChange Site-Directed Mutagenesis Kit from Agilent Technologies (Santa Clara, CA) to make the desired mutations as described by Luttik and coworkers [100], resulting in plasmid p416-TEF-sc$aro4_{k229l}$ (**Table 2.1**, at the end of the chapter).

5.2.2 Strain construction

All *S. cerevisiae* strains were constructed from BY4741 (*Mat a; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*) as the initial strain (**Table 2**, at the end of the chapter). Plasmids were transformed using the EZ Yeast Transformation II Kit from Zymo Research Corp. (Irvine, CA). Gene knockouts were generated using the "delete and repeat" method [123]. Gene disruption cassettes containing the *KanMX* selectable marker flanked by

loxP sites (obtained by PCR of the pUG6 plasmid [123]) were produced with 40 basepairs of homology on either side of each target integration site. Following yeast transformations, colonies were selected on 200mg/L G418 and PCR confirmed (see **Table 2.1** for primers, located at the end of the chapter).

Multiple gene knockouts were achieved using an interspersed step with Cre recombinase to excise the selection marker between the *loxP* sites in the disruption cassette. Cre recombinase was expressed using the inducible *GAL1* promoter on plasmid pSH47 [123]. Once marker removal was achieved, the strain was grown in YPD plus 1g/L 5-flouroorotic acid to encourage loss of the *URA3* containing pSH47 plasmid [169].

The mutant *aro4*$_{k229l}$ was integrated into the *ARO4* locus by cloning the expression cassette, including promoter and terminator, from p416-GPD- sc*aro4*$_{k229l}$ and inserting it into the pUG6 plasmid in order to create an integration cassette with the *KanMX* selectable marker. The insertion cassette was then cloned and transformed as described above for the gene disruption cassettes.

To integrate ECL_01944$_{opt}$, the expression cassette from p425-GPD-ECL_01944$_{opt}$, including the promoter and terminator, was cloned into vector pITy3 [101]. The resulting vector, pITy-GPD- ECL_01944$_{opt}$, was digested with the ScaI restriction enzyme to create a linear integration cassette for multiple integrations into the Ty2 δ sites [101]. Clones with multiple integrations were preferentially selected on higher concentration G418 plates (500 mg/L).

5.2.3 Enzyme activity assays

Dehydroshikimate (DHS) dehydratase and catechol 1,2-dioxygenase activities were assayed using total cell protein extract, which was obtained using the Pierce Y-PER Yeast Protein Extraction Reagent and EDTA-free Halt Protease Inhibitor from Thermo Fisher Scientific (Waltham, MA). The protein extraction was performed according to Y-PER reagent instructions. Protein concentration was determined using the BCA method with a Pierce BCA kit obtained from Thermo Fisher Scientific (Waltham, MA). Spectrophotometric assays for enzyme activity were then performed [170, 171]. For the DHS dehydratase, 300μg of protein was added to a cuvette containing excess Y-PER reagent and DHS such that the final concentration of DHS was 0.1-0.75 mM in a total volume of 1 mL. A reading of the absorbance at 290 nm was taken every 20 seconds for three minutes. For the catechol 1,2-dioxygenase, 300μL of protein (at approximately 1000 ug/mL) was added to a cuvette containing 100mM potassium phosphate buffer (pH=7.5) and catechol such that the final concentration of catechol was 0.1-0.4mM in a total volume of 1 mL. A reading of the absorbance at 288nm was taken every 2 seconds for approximately 90 seconds total. For both DHS dehydratase and catechol 1,2-dioxygenase, product formation was quantified using standard curves of the expected reaction product (PCA or cis,cis-muconic acid, respectively) in the presence of control protein extract (from cultures not producing the heterologous enzymes) as well as with varying amount of reactant (DHS or catechol, respectively) in order to take into account all possible contributions to absorbance. An Ultrospec 2100 pro UV/Visible

Spectrophotometer from Biochrom (Cambridge, UK) was used in the assays. DHS was obtained from Sigma-Aldrich (St. Louis, MO). Catechol, PCA, and cis,cis-muconic acid were obtained from Thermo Fisher Scientific (Waltham, MA).

5.2.4 Strain characterization

High pressure liquid chromatography (HPLC) was used to measure the production of muconic acid and pathway intermediates in *S. cerevisiae* cultures. Strains were pre-cultured in 5 mL aliquots for two days and used to inoculate 30 mL flask cultures at OD_{600}=0.25. After a designated time (usually 48 hours for initial characterizations), the OD_{600} was measured and a 1 mL sample was taken and pelleted for 5 min. at 3,000x g. The supernatant was filtered using a 0.2 micron syringe filter from Corning Incorporated (Corning, NY). Samples were then separated using a HPLC Ultimate 3000 from Dionex (Sunnyvale, CA) and a Zorbax SB-Aq column from Agilent Technologies (Santa Clara, CA). A 2.0 μL injection volume was used in a mobile phase composed of an 84:16 ratio of 25mM potassium phosphate buffer (pH=2.0) to acetonitrile with a flow rate of 1.0 mL/min. The column temperature was maintained at 30ºC and the UV-Vis absorption was measured at 280nm. Standards, including DHS, PCA, catechol, and cis,cis-muconic acid, were purchased from Sigma-Aldrich (St. Louis, MO). Cis, trans-muconic acid was generously provided by Draths Corporation.

Glucose utilization in the 30 mL flask cultures was measured using a YSI 7100 MBS from YSI Life Sciences (Yellow Springs, Ohio) according to manufacturer instructions. Culture supernatant was diluted 1:10 in DI water prior to measurement.

5.2.5 RT-PCR Analysis

The relative abundance of heterologous mRNA was determined using quantitative RT-PCR. RNA was extracted from mid-log phase cells using the Ambion Yeast Ribo-Pure Kit (Life Technologies, Carlsbad, CA) and cDNA was prepared using the Applied Biosystems High Capacity Reverse Transcription Kit (Life Technologies, Carlsbad, CA). Primers were designed using the Primer Quest utility from Integrated DNA Technologies (Coralville, Iowa), including primers for ECL_01944_{opt} (TGCATGGTTTCTCATTGTGACGGC and CAACATACAAACACTGGCCACGCT) and for ALG9 (ATCGTGAAATTGCAGGCAGCTTGG and CATGGCAACGGCAGAAGGCAATAA), which was used as the housekeeping gene. Quantitative PCR was performed on a ViiA7 Real Time PCR System (Life Technologies, Carlsbad, CA) using Fast Start SYBR Green Master Mix (Roche, Penzberg, Germany), following the manufacturer's instructions with an annealing temperature of 58°C.

5.2.6 Flux balance analysis calculations

Flux balance analysis calculations were performed on a Dell PC using MATLAB (Mathworks, Natick, MA) and the Cobra Toolbox add-in [18, 172]. The iMM904 genomic scale model was used [109]. The reactions representing the heterologous enzymatic activity of DHS dehydratase, PCA decarboxylase, and catechol 1,2-dioxygenase were added to the model as shown in Table 2.3. Maximum theoretical

yields were calculated by setting the muconic acid production reaction 'EX_MUA' as the objective function and solving the system of linear equations.

5.3 Materials and Methods for Chapter 3

5.3.1 Plasmid construction

UAS_{aro} elements were amplified from BY4741 gDNA (Wizard Kit, Zymo Research), purified, restriction digested and ligated generate p416-1x UAS_{aro} -LeuMin-yECitrine. Additional UAS_{aro} elements were sequentially added through restriction cloning. The 4x UAS_{aro} cassette and p416-CYC-yECitrine were PCR amplified, gel extracted and used to Gibson assemble p416-5x UAS_{aro} -CYC1-yECitrine [173]. The 5xUAS_{aro} cassette was then digested out with BamHI/PmeI and ligated into p416-HXT7-yECitrine, p416-CORE1-yECitrine, p416-LeuMin-yECitrine.

Yeast homologous recombination was used to assemble p415-TEF-aro80, by transforming the linearized fragments into BY4741 and selecting on CSM-LEU plates followed by purifying the resulting plasmids [174]. The S551I point mutation was generated using the Quickchange II kit (Strategene), while the T675S point mutation and p416-CORE1-yECitrine vector were constructed using inverse PCR followed by blunt end ligation. A table of resulting strains from plasmid transformations is provided in **Table 3.2**.

5.3.2 ARO80 Library Preparation

ARO80 gene was amplified from the p415-TEF-ARO80 plasmid DNA (primers listed in **Table 3.1**) using the Genemorph II kit (Strategene, La Jolla, Ca) according to

manufacturer's recommendations. This PCR product was confirmed by gel electrophoresis and digested overnight with SpeI-HF and XhoI. The p415-TEF vector was digested, phosphatased and insert ligated overnight and transformed into *E. coli* and resulting transformants spread over 15cm plates, reaching a library size of $4x10^4$. *E. coli* colonies were then scrapped from the plates, votexed, glycerol stocked and miniprepped (Thermo). This library was then transformed into the *aro80Δ*- p416-4x UAS_{aro}-LeuMin-yECitrine strain using the high-efficiency yeast transformation protocol [175] and resuspended in a 300ml liquid culture. A fraction was plated in order to estimate the effective library size present of $5x10^5$.

5.3.3 Flow Cytometry and FACS

The yeast cultures were inoculated in triplicate from glycerol stock in CSM-URA or CSM-URA-LEU and grown until stationary phase. All cultures were then inoculated at an OD_{600} of 0.01 in fresh media and grown to mid-log phase in 30℃ orbital shaker for 14-16 hours. Induction was accomplished by subculturing from stationary phase in CSM into media containing 500mg/L or 1g/L L-Tryptophan or 20g/L galactose. Fluorescence was analyzed on the Fortessa Flowcytometer (BD Biosciences) using the yECitrine fluorophore. 10,000 events were gathered at a flow rate of 1,000 events per second and analyzed using the FlowJo software suite. Data for the Digital-to-Analog converter described in **Figure 3.5** were grown in a 96-deep well block, 1ml culture volume and the fluorescence analyzed using the Accuri Flowcytometer (BD Biosciences). An average of

the fluorescence and standard deviation within biological replicates is provided. All experiments within a single graph were conducted on the same day at the same time.

The top 1% fluorescing cells from the aro80-library were sorted using the BD FACS Aria Cell sorter. Recovered cells were grown for 24 hours at 30℃in 2ml of CSM-URA-LEU media and plated to solid media. Individual colonies were randomly selected, inoculated to 2ml of CSM-URA-LEU and fluorescence compared to the control strains.

5.3.4 qPCR Analysis

S. cerevisiae BY4741 strains expressing amplifier circuit and GPD-YFP plasmids were inoculated from glycerol stock into CSM-URA-LEU and CSM-URA respectively. These were cultured for 48 hours and then inoculated into fresh media at OD_{600} of 0.01 in 2ml of fresh media and cultured for 15 hours with shaking. Total RNA was extracted from 1 OD unit of cells (Quick-RNA Miniprep, Zymo Research). 500ng of RNA was reverse transcribed (High Capacity cDNA Reverse Transcription Kit, Applied Biosystems) and quantified in triplicate (SYBR Green PCR Master Mix, Life Technologies) after RNA extraction. Transcript levels were quantified using primers previously designed to target yECitrine (TTCTGTCTCCGGTGAAGGTGAA and TAAGGTTGGCCATGGAACTGGCAA)[62] and that of the housekeeping gene ALG9 (ATCGTGAAATTGCAGGCAGCTTGG and CATGGCAACGGCAGAAGGCAATAA) [176]. Reactions were performed on the Viia 7 Real Time PCR Instrument (Life Technologies) and data analyzed using Viia 7 Software.

5.4 Materials and Methods for Chapter 4

5.4.1 Plasmid Construction

Gibson Assembly [173] was used to construct the PugM vector with ECL_01944opt under control of the UASCLB–UASCIT–UASTEF-GPD hybrid promoter [62] followed by the Tprm9 high capacity terminator [131] with parts amplified with primers listed in Table 1. The PugM and p425-TEF-pa5_5120opt/GPD-caHQD2opt vectors were used to amplify parts to construct p426-TEF-pa5_5120opt-Tcyc-GPD-ECL_01944opt-Tprm9 through Gibson Assembly. MoClo assembly was used to construct the PAD1 integration vector. A "Type 3" entry vector was built for scPAD1 using BsmbI golden gate assembly. The PAD1 integration vector, scPAD1-IV, was assembled from "part vectors" listed in Table 4.1 using BsaI golden gate assembly [177].

5.4.2 Growth Rate Analysis

Strains of interest were precultured for 3 days in the appropriate selective medium, and 1 μL of this precultured was used as an inoculum for a 250 μL culture in the same selective medium. Growth rate measurements were then obtained using a Bioscreen C (Growth Curves USA).

5.4.3 Flow Cytometry

For the initial biosensor assay, yeast cultures were inoculated in triplicate from glycerol stock in CSM-URA for two days at stationary phase. All cultures were then

inoculated at an OD_{600} of 0.01 in fresh media and grown to mid-log phase in 30℃ orbital shaker for 14-16 hours. Induction was accomplished by subculturing from stationary phase in CSM into media containing 500mg/L L-Tryptophan, 500mg/L L-Phenylalanine or 100mg/L Tyrosine. Fluorescence was analyzed on the Fortessa Flowcytometer (BD Biosciences) using the yECitrine fluorophore. 10,000 events were gathered at a flow rate of 1,000 events per second and analyzed using the FlowJo software suite.

For the integrated fluorescent biosensor assay, yeast cultures were inoculated in triplicate from glycerol stock in CSM-HIS in a 96-deep well block, 1ml culture volume and the fluorescence analyzed using the Accuri Flowcytometer (BD Biosciences). An average of the fluorescence and standard deviation within biological replicates is provided. All experiments within a single graph were conducted on the same day at the same time.

5.4.4 EMS Mutagenesis

The EMS mutagenesis procedures were performed following the protocol described by Winston [163]. An overnight culture was cultivated to OD_{600} of 1.0. Cells were then harvested, washed and suspended with 0.1M sodium phosphate buffer (pH 7). 30ul of EMS was added and incubated along with an unmutagenized control for 1 h at 30℃, with agitation. The cells were then washed twice with 5% sodium thiosulfate to eliminate residual EMS and the cells were allowed to recover in CSM-URA. This

procedure was repeated on a portion of mutagenized cells. Kill fraction was calculated by plating a fraction from mutagenized and control populations.

5.4.5 Subculturing Procedure

Following EMS mutagenesis the recovered cells were sequentially subcultured in 30ml of CSM-URA media. After reaching stationary phase, 100ul of cells were transferred to 30ml of fresh selective media with increasing concentrations of G418 and analog. During this first round of selection the populations were subcultured six times and cultured for over 750 hours after which the unmutagenized control populations were unable to grow in the selective media conditions while the mutagenized populations continued to grow in 1000 mg/L G418 and 2mM 4-FP.

During the second round of selection, following recovery from EMS, the cultures were inoculated into 1000 mg/L G418 and 1mM 4-FP. Over the course of selection, the 4FP selective concentration was increased to 7.5mM which resulted in an observed growth rate difference between the populations. Following selection, colonies were plated and isolated. Colonies were selected and the strain was grown in YPD plus 1 g/L 5-flouroorotic acid to encourage loss of the URA3 containing pSH47 plasmid [169].

5.4.6 Tyrosine Quantification

Tyrosine quantification was confirmed using four different protocols. The first protocol used to quantify the data reported in **Figure 4.4** used cultures directly selected from CSM-URA plates following selection. OD_{600} measurements were taken and 100ul

of supernatant were isolated and assayed using nitrosonapthol chemical derivation using nitrosonapthol derivatization which had been previously developed [164] and analyzed using the BioTek Cytation 3 plate reader. Tyrosine concentration was calculated based on standards and production.

The second protocol used to quantify the data reported in **Figure 4.6** used cultures directly selected from YPD plates into 4ml of YPD liquid media. OD_{600} measurements were taken and cells were resuspended to reach 0.5 OD_{600} in 4ml of fresh 2% glucose CSM. After 48 hour of culturing, 300ul of supernatant extracted and quantified with the derivatization protocol listed above. The third protocol used to quantify the data reported in **Figure 4.7** used cultures directly selected glycerol storage into 4ml of 4% glucose CSM. OD_{600} measurements were taken and cells were resuspended to reach 0.1 OD_{600} in 30ml of fresh 4% glucose CSM in 250ml shake flasks. After 70 hour of culturing, 300ul of supernatant extracted and quantified with the derivatization protocol listed above.

The fourth protocol was used to quantify the data reported in **Figure 4.8.** 96-deep well blocks were inoculated from glycerol stock and grown to stationary phase. OD_{600} was measured with BioTek Cytation 3 plate reader and the block spun down and cells washed with water. The cells were then cultured overnight in minimal media without YNB. After 16 hours of incubation, the cells were pelleted and the supernatant extracted. Tyrosine in the supernatant was quantified using nitrosonapthol derivatization described above. Tyrosine concentration was calculated based on standards and production normalized per OD unit.

5.4.7 HPLC

High performance liquid chromatography (HPLC) was used to measure the production of muconic acid and pathway intermediates in *S. cerevisiae* cultures. Strains were pre-cultured in 4mL aliquots until they reached stationary phase and used to inoculate 30 mL flask cultures at OD_{600} = 0.1. At designated times (72, 120 and 168 h), the OD_{600} was measured and a 1.5 mL sample was taken and pelleted for 5 min. at 3000 x g. The supernatant was filtered using a nylon 0.2 mm syringe filter from Corning Incorporated.

Samples were then separated using a HPLC Ultimate 3000 from Dionex with dilution as required. Composite pathway products were quantified using the Zorbax SB-Aq column from Agilent Technologies. A 10.0 uL injection volume was used in a mobile phase composed of an 84:16 ratio of ddH2O with 0.1% Trifluoroacetic Acid to acetonitrile with 0.1% TFA with a flow rate of 1.0 mL/min. Column temperature was maintained at 30℃ and UV–Vis absorption was measured at 280 nm. Peaks were compared to a standard curve including PCA, catechol, and cis,cis-muconic acid, purchased from Sigma-Aldrich. Cis, trans-muconic acid had been previously provided by Draths Corporation. Reported value represents the highest titer reached during the three timepoints assayed.

Glucose quantification was accomplished using the Aminex HPX-87p Carbohydrate Column from Bio-Rad Laboratories with RI detection performed by

Refractomax 520 modular unit. The mobile phase was 100% ddH2O, with a flow rate of 0.6mL/min. The column temperature was maintained at 85℃.

5.4.8 Bioreactor Fermentations

Bioreactor fermentations were run using a Bioflo 115 (New Brunswick) using CSM-LUHW with 4% glucose as a fed-batch process with up to three media spikes throughout the run. All fermentations were inoculated to an initial OD_{600} = 0.2 in 1.7L of media. Dissolved oxygen was maintained through cascaded agitation with a constant air input flow rate of 0.5 SLPM. The pH was maintained at 5.0 with 1M NaOH and temperature was maintained at 30℃. Throughout the run, 21 mL of samples were removed from the reactor and fermentations lasted up to 12 days.

Chapter 6: Conclusions and Major Findings

6.1 Major Findings

The experiments described here represent a synthesis of metabolic engineering, synthetic biology and adaptive laboratory evolution. We begin with the initial work of developing a composite pathway capable of producing muconic acid in yeast followed up by rational metabolic engineering to improve titers. To further strain development, we engineer a biosensor capable of detecting the downstream products of the shikimate pathway, aromatic amino acids. We demonstrate the utility of this biosensor through the development of a mutant ARO80p as a proof of concept. Finally, we use this biosensor to develop a strain of yeast capable of increased muconic acid production titers which represents the highest titers produced of this molecule in a yeast host.

For our first aim, we screened seventeen potential enzymes to build a composite pathway capable of producing muconic acid in *S. cerevisiae* utilizing enzymatic activity assays to confirm enzyme function. This pathway consisted of: ca*HQD2*$_{opt}$, pa5_5120$_{opt}$ and kpAroY$_{opt}$. With a functional muconic acid pathway drawing off of flux from the shikimate pathway built, we removed the feedback inhibition previously shown to exist within the aromatic amino acid biosynthetic pathway by knocking-out ARO3/ARO4 and integrating the feedback resistant mutant of *aro4*$_{k229l}$. However, this composite pathway needed balancing as we were producing a high concentration of PCA relative to our muconic acid production. Thus, we optimized our pathway by employing a higher performing PCA decarboxylase, the AroY gene ECL_01944opt, and transferred the pathway to a combination of high copy plasmids and genomic integration. Next, we used flux balance analysis to predict metabolic changes which could re-route flux into our composite pathway and improve muconic acid yields and titers. The solution predicted

for increasing flux into the starting point of the shikimate pathway, erythrose-4-phosphate and phosphoenolpyruvate was to overexpress the transketolase gene TKL1 and subsequently knockout the glucose-6-phosphate dehydrogenase gene, ZWF1, to force entry into the pentose phosphate pathway to occur via transketolase. With final media optimization, we were able to report a titer of 141 mg/L muconic acid, 24 times the value of the initial strain.

The second goal was to develop an aromatic amino acid biosensor and utilize this sensor for the engineering of its activator, the transcription factor ARO80p. This application would function as a proof of concept for further applications of this biosensor for engineering *S. cerevisiae* at the protein or genome level for higher production of aromatic amino acids and muconic acid. We used this biosensor to accomplish the engineering of both cis DNA elements and the trans-acting factor responsible for their activation. The first step in this process was to develop a hybrid promoter based on the ARO9 promoter previously demonstrated to be sensitive to exogenous concentrations of aromatic amino acids. We cloned this UASaro element 5' of the LeuMin minimal promoter element and demonstrated that it could be induced as well as the wild-type ARO9 promoter.

Having demonstrated that we could employ a hybrid approach to developing an ARO9 based biosensor, we expanded hybrid promoter to include 4x-UASaro elements and demonstrated that in conjunction with ARO80wt overexpression, that we could observe a stronger signal through the hybrid promoter than the native ARO9 promoter while maintaining a 2.5 fold induciblity with exogenous tryptophan. Having developed a cis-architecture capable of stronger expression, we turned our attention to increasing the strength of its trans-acting factor ARO80p. We then screened a protein library of ARO80

and isolated the ARO80mut containing two point mutations I551S and S675T. When paired with the 4xUASaro-LeuMin promoter, it resulted in constitutive expression 5-fold of the ARO80wt, while still maintaining its inducibility.

We then employed this ARO80mut to produce two synthetic genetic circuits. The first circuit implemented was that of an "amplifier" to generate ultra-strong promoters. The ARO80mut was expressed under control of the strong TEF promoter, while 4 or 5xUASaro elements were placed 5' of full length promoters HXT7 and CYC1 as well as LeuMin and CORE1 minimal promoters. At best, this resulted in 15-fold amplification compared to the core promoters and up to a 2-fold increase in protein expression compared to GPD, one of the strongest available constitutive promoters.

We then implemented a circuit capable of staged, multi-output response, previously termed a "Digital-to-Analog converter". We placed the aro80mut under the GAL1 promoter while using it to drive expression from three hybrid promoters constructs. By controlling the concentration of tryptophan and use of glucose or galactose, we were able to produce staged expression outputs. We then were able to remove one of the stages by knocking out the endogenous ARO80wt, limiting the tryptophan induction without growth in galactose. The application of the biosensor to engineer the transcription factor responsible for its activation provides a useful proof-of-concept for further engineering efforts.

The third goal was to utilize this biosensor to develop improved yeast strains for muconic acid production. To investigate potential mutations beneficial for muconic acid production, we employed an adaptive laboratory evolution experiment (ALE) with the aromatic amino acid production functioning as a surrogate for muconic acid, as aromatic amino acids are derived from the same pathway. We first demonstrated that this ARO9

based biosensor was capable of detecting intracellular concentrations of aromatic amino acids as well as extracellular by expressing it in our previously engineered strain with *aro3Δ; aro4Δ::*P_{GPD}*-aro4*$_{k229l}$*; zwf1Δ* (ENG) and comparing it to expression in BY4741. We then demonstrated that the biosensor could detect further improvements by inducing the biosensor in BY4741 and ENG with exogenous tryptophan, tyrosine and phenylalanine. Next, we turned to developing a growth based selection scheme for use in ALE. We built an aromatic amino acid inducible G418 antibiotic resistance vector using the 4xUASaro-LeuMin promoter to drive expression of the weak KanNeo gene.

We then mutagenized the ENG strain expressing the KanNeo biosensor with EMS and then performed serial sub-culture to enrich the population of cells for improved growth under increasing selective conditions of G418 and the anti-metabolites 4-Fluorophenylalanine to provide additional selection pressures for producing aromatic amino acids, This resulted in three ALE isolated strains which we showed, using tyrosine quantification and a fluorescent biosensor, to have improved aromatic amino acid production. These were then used for a second round of ALE resulting in four strains with improved aromatic amino acid production. The muconic acid composite pathway was transformed into these strains and they showed a three-fold improvement in pathway output relative to the ENG strain. Next we rerouted flux into our composite pathway through a truncated ARO1 protein and overexpression of the native yeast PCA decarboxylase PAD1. This resulted in a strain of yeast capable of 550mg/L muconic acid production in flask. We then scaled this up to 3 L bioreactor fermentations and employed dissolved oxygen control to result in 1.94 g/L production, the highest reported titer of muconic acid produced in *S. cerevisiae* as well as the highest reported titer of any shikimate derivative.

In summary, this chapter represents the first demonstration of muconic acid production in *S. cerevisiae* and the application of rational and evolutionary techniques to bring it from a proof of concept (mg/L) to industrially relevant levels of production (g/L). Contributing to this work was the development and application of a biosensor capable of detecting intracellular and extracellular aromatic amino acids and its use in an adaptive laboratory evolution scheme. This biosensor was then utilized in a cis-trans combined engineering approach to develop an ultra-strong expression system for use in *S. cerevisiae* resulting in a promoter twice as strong as GPD, one of the strongest endogenous promoters in yeast.

6.2 Proposals for Future Work

The studies described here demonstrate the combination of metabolic engineering with evolutionary studies. This could be further expanded by using this aromatic amino acid biosensor in protein engineering applications to engineer improved TKL1, ARO1 other regulatory enzymes such as GCN4 which are involved in aromatic amino acid metabolism. The availability of biosensors for use in *S. cerevisiae* is expanding, and the ALE scheme demonstrated could be easily retuned for improved flux through other metabolites.

The ARO80mut trans-acting factor facilitating a strong output from small and modular promoters it could be used to enable the expression of other complex genetic circuits. With the small DNA footprint of the UASaro element, it could be used to express heterologous promoters from bacterial species to facilitate the importing of bacterial operons into eukaryotic hosts. Finally, other researchers have shown that ARO80 expression can be modulated by temperature, presumably by allowing more

aromatic amino acids to enter the cell. This could be used to provide a further modulator on the ultra-strong promoters capable of staged outputs that we initially demonstrated.

Finally, it would be very beneficial to sequence the final strains isolated through the ALE process. Through sequencing, we would be able to determine mechanisms for increased flux through the shikimate pathway which are common to all of them and unique to our final production strain. This could inform future work for the production of aromatic compounds in general, as well as generating targets for protein engineering to further increase muconic acid titers. It is of special interest that our final strain was able to gain a significant improvement from ARO1t relative to other ALE strains, suggesting that it is complementing another mutation.

Chapter 7: References

[1] Curran, K. a., Leavitt, J. M., Karim, A. S., Alper, H. S., Metabolic engineering of muconic acid production in *Saccharomyces cerevisiae*. *Metabolic engineering* 2013, *15*, 55-66.
[2] Leavitt, J. M., Tong, A., Tong, J., Pattie, J., Alper, H. S., Coordinated transcription factor and promoter engineering to establish strong expression elements in *Saccharomyces cerevisiae*. *Biotechnology journal* 2016, n/a-n/a.
[3] Dwyer, J., The century of biology: three views. *Sustainability Science* 2008, *3*, 283-285.
[4] Curran, K. a., Alper, H. S., Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metabolic engineering* 2012, *14*, 289-297.
[5] Sun, J., Alper, H. S., Metabolic engineering of strains: from industrial-scale to lab-scale chemical production. *Journal of industrial microbiology & biotechnology* 2015, *42*, 423-436.
[6] Van Dien, S., From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. *Current opinion in biotechnology* 2013, *24*, 1061-1068.
[7] Liu, L., Redden, H., Alper, H. S., Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. *Current opinion in biotechnology* 2013.
[8] Nielsen, J., Yeast cell factories on the horizon. *Science* 2015, *349*, 1050-1051.
[9] Lee, J. W., Kim, H. U., Choi, S., Yi, J., Lee, S. Y., Microbial production of building block chemicals and polymers. *Current opinion in biotechnology* 2011, *22*, 758-767.
[10] Nielsen, J., Larsson, C., van Maris, A., Pronk, J., Metabolic engineering of yeast for production of fuels and chemicals. *Current opinion in biotechnology* 2013, *24*, 398-404.
[11] Hong, K.-K., Nielsen, J., Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cellular and Molecular Life Sciences* 2012, *69*, 2671-2690.
[12] Liu, L., Pan, A., Spofford, C., Zhou, N., Alper, H. S., An evolutionary metabolic engineering approach for enhancing lipogenesis in *Yarrowia lipolytica*. *Metabolic engineering* 2015, *29*, 36-45.
[13] Wriessnegger, T., Pichler, H., Yeast metabolic engineering – Targeting sterol metabolism and terpenoid formation. *Progress in Lipid Research* 2013, *52*, 277-293.
[14] McKenna, R., Thompson, B., Pugh, S., Nielsen, D. R., Rational and combinatorial approaches to engineering styrene production by *Saccharomyces cerevisiae*. *Microb Cell Fact* 2014, *13*.
[15] Xie, N.-Z., Liang, H., Huang, R.-B., Xu, P., Biotechnological production of muconic acid: current status and future prospects. *Biotechnology Advances* 2014, *32*, 615-622.
[16] Yadav, V. G., De Mey, M., Giaw Lim, C., Kumaran Ajikumar, P., Stephanopoulos, G., The future of metabolic engineering and synthetic biology: Towards a systematic practice. *Metabolic engineering* 2012, *14*, 233-241.

[17] Horwitz, Andrew A., Walter, Jessica M., Schubert, Max G., Kung, Stephanie H.*, et al.*, Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas. *Cell Systems* 2015, *1*, 88-96.
[18] Becker, S. A., Feist, A. M., Mo, M. L., Hannum, G.*, et al.*, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature protocols* 2007, *2*, 727-738.
[19] Österlund, T., Nookaew, I., Nielsen, J., Fifteen years of large scale metabolic modeling of yeast: Developments and impacts. *Biotechnology Advances* 2012, *30*, 979-988.
[20] Fletcher, E., Krivoruchko, A., Nielsen, J., Industrial systems biology and its impact on synthetic biology of yeast cell factories. *Biotechnol Bioeng* 2016, *113*, 1164-1170.
[21] Mahr, R., Frunzke, J., Transcription factor-based biosensors in biotechnology: current state and future prospects. *Applied microbiology and biotechnology* 2016, *100*, 79-90.
[22] Portnoy, V. a., Bezdan, D., Zengler, K., Adaptive laboratory evolution--harnessing the power of biology for metabolic engineering. *Current opinion in biotechnology* 2011, *22*, 590-594.
[23] Williams, T. C., Pretorius, I. S., Paulsen, I. T., Synthetic Evolution of Metabolic Productivity Using Biosensors. *Trends in biotechnology* 2016, *34*, 371-381.
[24] Suastegui, M., Matthiesen, J. E., Carraher, J. M., Hernandez, N.*, et al.*, Combining Metabolic Engineering and Electrocatalysis: Application to the Production of Polyamides from Sugar. *Angewandte Chemie International Edition* 2016, *55*, 2368-2373.
[25] Bui, V., Lau, M. K., MacRae, D., Schweitzer, D., Google Patents 2013.
[26] Draths, K. M., Frost, J. W., Environmentally compatible synthesis of adipic acid from D-glucose. *Journal of the American Chemical Society* 1994, *116*, 399-400.
[27] In, M.-J., Kim, D. C., Chae, H. J., Downstream process for the production of yeast extract using brewer's yeast cells. *Biotechnology and Bioprocess Engineering* 2005, *10*, 85-90.
[28] Jeffries, T. W., Engineering yeasts for xylose metabolism. *Current opinion in biotechnology* 2006, *17*, 320-326.
[29] Hansen, E. H., Moller, B. L., Kock, G. R., Bunner, C. M.*, et al.*, De novo biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). *Appl Environ Microbiol* 2009, *75*, 2765-2774.
[30] Krömer, J. O., Nunez-Bernal, D., Averesch, N. J. H., Hampe, J.*, et al.*, Production of aromatics in *Saccharomyces cerevisiae*—A feasibility study. *Journal of biotechnology* 2013, *163*, 184-193.
[31] Vannelli, T., Wei Qi, W., Sweigard, J., Gatenby, A. A., Sariaslani, F. S., Production of p-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metabolic engineering* 2007, *9*, 142-151.
[32] Wang, Y., Yu, O., Synthetic scaffolds increased resveratrol biosynthesis in engineered yeast cells. *Journal of biotechnology* 2012, *157*, 258-260.

[33] Jiang, H., Wood, K. V., Morgan, J. A., Metabolic Engineering of the Phenylpropanoid Pathway in *Saccharomyces cerevisiae*. *Applied and environmental microbiology* 2005, *71*, 2962-2969.
[34] Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E.*, et al.*, Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metabolic engineering* 2012, *14*, 104-111.
[35] Almario, M. P., Reyes, L. H., Kao, K. C., Evolutionary engineering of *Saccharomyces cerevisiae* for enhanced tolerance to hydrolysates of lignocellulosic biomass. *Biotechnology and bioengineering* 2013, *110*, 2616-2623.
[36] Gu, H., Zhang, J., Bao, J., Inhibitor analysis and adaptive evolution of *Saccharomyces cerevisiae* for simultaneous saccharification and ethanol fermentation from industrial waste corncob residues. *Bioresource Technology* 2014, *157*, 6-13.
[37] Jiménez, J., Benítez, T., Adaptation of Yeast Cell Membranes to Ethanol. *Applied and environmental microbiology* 1987, *53*, 1196-1198.
[38] Kildegaard, K. R., Hallström, B. M., Blicher, T. H., Sonnenschein, N.*, et al.*, Evolution reveals a glutathione-dependent mechanism of 3-hydroxypropionic acid tolerance. *Metabolic engineering* 2014, *26*, 57-66.
[39] Lee, D.-H., Palsson, B. Ø., Adaptive Evolution of *Escherichia coli* K-12 MG1655 during Growth on a Nonnative Carbon Source, l-1,2-Propanediol. *Applied and environmental microbiology* 2010, *76*, 4158-4168.
[40] Lee, J.-Y., Seo, J., Kim, E.-S., Lee, H.-S., Kim, P., Adaptive evolution of *Corynebacterium glutamicum* resistant to oxidative stress and its global gene expression profiling. *Biotechnol Lett* 2013, *35*, 709-717.
[41] Oide, S., Gunji, W., Moteki, Y., Yamamoto, S.*, et al.*, Thermal and Solvent Stress Cross-Tolerance Conferred to *Corynebacterium glutamicum* by Adaptive Laboratory Evolution. *Applied and environmental microbiology* 2015, *81*, 2284-2298.
[42] Hu, H., Wood, T. K., An evolved *Escherichia coli* strain for producing hydrogen and ethanol from glycerol. *Biochemical and biophysical research communications* 2010, *391*, 1033-1038.
[43] Lee, S.-W., Oh, M.-K., A synthetic suicide riboswitch for the high-throughput screening of metabolite production in *Saccharomyces cerevisiae*. *Metabolic engineering* 2015, *28*, 143-150.
[44] Mahr, R., Gätgens, C., Gätgens, J., Polen, T.*, et al.*, Biosensor-driven adaptive laboratory evolution of l-valine production in *Corynebacterium glutamicum*. *Metabolic engineering* 2015, *32*, 184-194.
[45] Eggeling, L., Bott, M., Marienhagen, J., Novel screening methods — biosensors. *Current opinion in biotechnology* 2015, *35*, 30-36.
[46] Tang, S.-Y., Qian, S., Akinterinwa, O., Frei, C. S.*, et al.*, Screening for Enhanced Triacetic Acid Lactone Production by Recombinant *Escherichia coli* Expressing a Designed Triacetic Acid Lactone Reporter. *Journal of the American Chemical Society* 2013, *135*, 10099-10103.

[47] Iraqui, I., Vissers, S., Cartiaux, M., Urrestarazu, a., Characterisation of *Saccharomyces cerevisiae* ARO8 and ARO9 genes encoding aromatic aminotransferases I and II reveals a new aminotransferase subfamily. *Molecular & general genetics : MGG* 1998, *257*, 238-248.
[48] Iraqui, I., Vissers, S., André, B., Urrestarazu, a., Transcriptional induction by aromatic amino acids in *Saccharomyces cerevisiae*. *Molecular and cellular biology* 1999, *19*, 3360-3371.
[49] Kennell, D., Riezman, H., Transcription and translation initiation frequencies of the *Escherichia coli* lac operon. *Journal of molecular biology* 1977, *114*, 1-21.
[50] Da Silva, N. a., Srikrishnan, S., Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. *FEMS yeast research* 2012, *12*, 197-214.
[51] Craven, S. H., Ezezika, O. C., Haddad, S., Hall, R. A.*, et al.*, Inducer responses of BenM, a LysR-type transcriptional regulator from *Acinetobacter baylyi* ADP1. *Molecular Microbiology* 2009, *72*, 881-894.
[52] Quandt, E. M., Hammerling, M. J., Summers, R. M., Otoupal, P. B.*, et al.*, Decaffeination and Measurement of Caffeine Content by Addicted *Escherichia coli* with a Refactored N-Demethylation Operon from *Pseudomonas putida* CBB5. *ACS synthetic biology* 2013, *2*, 301-307.
[53] Frenzel, A., Hust, M., Schirrmann, T., Expression of recombinant antibodies. *Frontiers in immunology* 2013, *4*, 217-217.
[54] Overton, T. W., Recombinant protein production in bacterial hosts. *Drug discovery today* 2014, *19*, 590-601.
[55] Frasch, H.-J., Medema, M. H., Takano, E., Breitling, R., Design-based re-engineering of biosynthetic gene clusters: plug-and-play in practice. *Current opinion in biotechnology* 2013, *24*, 1144-1150.
[56] Knight, T., Idempotent Vector Design for Standard Assembly of Biobricks. *MIT Synthetic Biology Working Group Technical Reports* 2003.
[57] Galdzicki, M., Clancy, K. P., Oberortner, E., Pocock, M.*, et al.*, The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nature biotechnology* 2014, *32*, 545-550.
[58] Salis, H. M., Mirsky, E. a., Voigt, C. a., Automated design of synthetic ribosome binding sites to control protein expression. *Nature biotechnology* 2009, *27*, 946-950.
[59] Woo, S., Yang, J.-s., Kim, I., Yang, J.*, et al.*, Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency. *Metabolic engineering* 2013, *15*, 67-74.
[60] Na, D., Lee, D., RBSDesigner: software for designing synthetic ribosome binding sites that yields a desired level of protein expression. *Bioinformatics (Oxford, England)* 2010, *26*, 2633-2634.
[61] Lanza, A. M., Curran, K. a., Rey, L. G., Alper, H. S., A condition-specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae*. *BMC systems biology* 2014, *8*, 33-33.

[62] Blazeck, J., Garg, R., Reed, B., Alper, H. S., Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Biotechnology and bioengineering* 2012, *109*, 2884-2895.
[63] Perez-Pinera, P., Ousterout, D. G., Brunger, J. M., Farin, A. M.*, et al.*, Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nature methods* 2013, *10*, 239-242.
[64] Alper, H., Fischer, C., Nevoigt, E., Stephanopoulos, G., Tuning genetic control through promoter engineering. *Proceedings of the National Academy of Sciences of the United States of America* 2005, *102*, 12678-12683.
[65] Liang, J., Ning, J. C., Zhao, H., Coordinated induction of multi-gene pathways in *Saccharomyces cerevisiae*. *Nucleic acids research* 2013, *41*, e54-e54.
[66] Blazeck, J., Liu, L., Redden, H., Alper, H., Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. *Applied and environmental microbiology* 2011, *77*, 7905-7914.
[67] Vogl, T., Ruth, C., Pitzer, J., Kickenweiz, T., Glieder, A., Synthetic core promoters for *Pichia pastoris*. *ACS synthetic biology* 2014, *3*, 188-191.
[68] Temme, K., Hill, R., Segall-Shapiro, T. H., Moser, F., Voigt, C. a., Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic acids research* 2012, *40*, 8773-8781.
[69] Munchel, S. E., Shultzaberger, R. K., Takizawa, N., Weis, K., Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay. *Molecular biology of the cell* 2011, *22*, 2787-2795.
[70] Farzadfard, F., Perli, S. D., Lu, T. K., Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS synthetic biology* 2013, *2*, 604-613.
[71] Lee, J. W., Na, D., Park, J. M., Lee, J.*, et al.*, Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat Chem Biol* 2012, *8*, 536-546.
[72] Lee, J. W., Kim, T. Y., Jang, Y.-S., Choi, S., Lee, S. Y., Systems metabolic engineering for chemicals and materials. *Trends in biotechnology* 2011, *29*, 370-378.
[73] Jang, Y.-S., Kim, B., Shin, J. H., Choi, Y. J.*, et al.*, Bio-based production of C2–C6 platform chemicals. *Biotechnology and bioengineering* 2012, *109*, 2437-2459.
[74] de Jong, B., Siewers, V., Nielsen, J., Systems biology of yeast: enabling technology for development of cell factories for production of advanced biofuels. *Current opinion in biotechnology* 2012, *23*, 624-630.
[75] Zhou, S., Yomano, L. P., Shanmugam, K. T., Ingram, L. O., Fermentation of 10% (w/v) Sugar to D(−)-Lactate by Engineered Escherichia coli B. *Biotechnol Lett* 2005, *27*, 1891-1896.
[76] Zhou, Q., Shi, Z.-Y., Meng, D.-C., Wu, Q.*, et al.*, Production of 3-hydroxypropionate homopolymer and poly(3-hydroxypropionate-co-4-hydroxybutyrate) copolymer by recombinant *Escherichia coli*. *Metabolic engineering* 2011, *13*, 777-785.
[77] Yumoto, I., Ikeda, K., Direct fermentation of starch to L-(+)-lactic acid using *Lactobacillus amylophilus*. *Biotechnol Lett* 1995, *17*, 543-546.

[78] Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C.*, et al.*, Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol* 2011, *7*, 445-452.
[79] Stols, L., Donnelly, M. I., Production of succinic acid through overexpression of NAD(+)-dependent malic enzyme in an Escherichia coli mutant. *Applied and environmental microbiology* 1997, *63*, 2695-2701.
[80] McKenna, R., Nielsen, D. R., Styrene biosynthesis from glucose by engineered E. coli. *Metabolic engineering* 2011, *13*, 544-554.
[81] Li, Z.-J., Shi, Z.-Y., Jian, J., Guo, Y.-Y.*, et al.*, Production of poly(3-hydroxybutyrate-co-4-hydroxybutyrate) from unrelated carbon sources by metabolically engineered *Escherichia coli*. *Metabolic engineering* 2010, *12*, 352-359.
[82] Ikushima, S., Fujii, T., Kobayashi, O., Yoshida, S., Yoshida, A., Genetic Engineering of Candida utilis Yeast for Efficient Production of L-Lactic Acid. *Bioscience, Biotechnology, and Biochemistry* 2009, *73*, 1818-1824.
[83] Antoniewicz, M. R., Kraynie, D. F., Laffend, L. A., González-Lergier, J.*, et al.*, Metabolic flux analysis in a nonstationary system: Fed-batch fermentation of a high yielding strain of *E. coli* producing 1,3-propanediol. *Metabolic engineering* 2007, *9*, 277-292.
[84] Burridge, E., Adipic acid. . *ICIS Chem. Bus.* 2011, *279* 43-43.
[85] Mirasol, F., PTA. . *ICIS Chem. Bus.* 2011, *279* 43-43.
[86] Tsai, S.-C., Tsai, L.-D., Li, Y.-K., An Isolated *Candida albicans* TL3 Capable of Degrading Phenol at Large Concentration. *Bioscience, Biotechnology, and Biochemistry* 2005, *69*, 2358-2367.
[87] Warhurst, A. M., Clarke, K. F., Hill, R. A., Holt, R. A., Fewson, C. A., Production of catechols and muconic acids from various aromatics by the styrene-degrader *Rhodococcus rhodochrous* NCIMB 13259. *Biotechnol Lett* 1994, *16*, 513-516.
[88] Wu, C.-M., Lee, T.-H., Lee, S.-N., Lee, Y.-A., Wu, J.-Y., Microbial synthesis of cis,cis-muconic acid by *Sphingobacterium* sp. GCG generated from effluent of a styrene monomer (SM) production plant. *Enzyme and Microbial Technology* 2004, *35*, 598-604.
[89] Neidle, E. L., Hartnett, C., Bonitz, S., Ornston, L. N., DNA sequence of the *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase I structural gene catA: evidence for evolutionary divergence of intradiol dioxygenases by acquisition of DNA sequence repetitions. *Journal of Bacteriology* 1988, *170*, 4874-4880.
[90] Niu, W., Draths, K. M., Frost, J. W., Benzene-free synthesis of adipic acid. *Biotechnology progress* 2002, *18*, 201-211.
[91] Naesby, M., Nielsen, S. V., Nielsen, C. A., Green, T.*, et al.*, Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in *Saccharomyces cerevisiae*. *Microb Cell Fact* 2009, *8*, 45.
[92] Qi, W. W., Vannelli, T., Breinig, S., Ben-Bassat, A.*, et al.*, Functional expression of prokaryotic and eukaryotic genes in *Escherichia coli* for conversion of glucose to p-hydroxystyrene. *Metab Eng* 2007, *9*, 268-276.

[93] Wang, Y., Halls, C., Zhang, J., Matsuno, M.*, et al.*, Stepwise increase of resveratrol biosynthesis in yeast *Saccharomyces cerevisiae* by metabolic engineering. *Metabolic engineering* 2011, *13*, 455-463.
[94] Hawkins, A. R., Francisco, A. J., Roberts, C. F., Cloning and characterization of the three enzyme structural genes QUTB, QUTC and QUTE from the quinic acid utilization gene cluster in *Aspergillus nidulans*. *Current Genetics* 1985, *9*, 305-311.
[95] Tsai, S. C., Li, Y. K., Purification and characterization of a catechol 1,2-dioxygenase from a phenol degrading *Candida albicans* TL3. *Archives of microbiology* 2007, *187*, 199-206.
[96] Yoshida, T., Inami, Y., Matsui, T., Nagasawa, T., Regioselective carboxylation of catechol by 3,4-dihydroxybenzoate decarboxylase of *Enterobacter cloacae* P. *Biotechnol Lett* 2010, *32*, 701-705.
[97] Ren, Y., Ren, Y., Zhou, Z., Guo, X.*, et al.*, Complete Genome Sequence of *Enterobacter cloacae* subsp. cloacae Type Strain ATCC 13047. *Journal of Bacteriology* 2010, *192*, 2463-2464.
[98] Mukai, N., Masaki, K., Fujii, T., Kawamukai, M., Iefuji, H., PAD1 and FDC1 are essential for the decarboxilation of phenylacrylic acid in *Saccharomyces cerevisiae*. *J Biosci Bioeng* 2010, *109*.
[99] Braus, G. H., Aromatic amino acid biosynthesis in the yeast *Saccharomyces cerevisiae*: a model system for the regulation of a eukaryotic biosynthetic pathway. *Microbiol Rev* 1991, *55*, 349-370.
[100] Luttik, M. A. H., Vuralhan, Z., Suir, E., Braus, G. H.*, et al.*, Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: quantification of metabolic impact. *Metab Eng* 2008, *10*, 141-153.
[101] Parekh, R. N., Shaw, M. R., Wittrup, K. D., An Integrating Vector for Tunable, High Copy, Stable Integration into the Dispersed Ty δ Sites of *Saccharomyces cerevisiae*. *Biotechnology progress* 1996, *12*, 16-21.
[102] Brochado, A. R., Matos, C., Møller, B. L., Hansen, J.*, et al.*, Improved vanillin production in baker's yeast through *in silico* design. *Microbial cell factories* 2010, *9*, 84-84.
[103] Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U., Lee, S. Y., Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Molecular systems biology* 2007, *3*, 149.
[104] Park, J. H., Lee, K. H., Kim, T. Y., Lee, S. Y., Metabolic engineering of *Escherichia coli* for the production of l-valine based on transcriptome analysis and in silico gene knockout simulation. *Proceedings of the National Academy of Sciences of the United States of America* 2007, *104*, 7797-7802.
[105] Becker, S. A., Feist, A. M., Mo, M. L., Hannum, G.*, et al.*, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protocols* 2007, *2*, 727-738.

[106] Asadollahi, M. A., Maury, J., Patil, K. R., Schalk, M.*, et al.*, Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through *in silico* driven metabolic engineering. *Metabolic engineering* 2009, *11*, 328-334.
[107] Alper, H., Jin, Y.-S., Moxley, J. F., Stephanopoulos, G., Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metabolic engineering* 2005, *7*, 155-164.
[108] Hong, S. H., Park, S. J., Moon, S. Y., Park, J. P., Lee, S. Y., In silico prediction and validation of the importance of the Entner–Doudoroff pathway in poly(3-hydroxybutyrate) production by metabolically engineered *Escherichia coli*. *Biotechnology and bioengineering* 2003, *83*, 854-863.
[109] Mo, M. L., Palsson, B. Ø., Herrgård, M. J., Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC systems biology* 2009, *3*, 1-17.
[110] Sprenger, G. A., Schörken, U., Sprenger, G., Sahm, H., Transketolase a of *Escherichia coli* K12. *European Journal of Biochemistry* 1995, *230*, 525-532.
[111] Grant, D. J., Patel, J. C., The non-oxidative decarboxylation of p-hydroxybenzoic acid, gentisic acid, protocatechuic acid and gallic acid by *Klebsiella aerogenes* (Aerobacter aerogenes). *Antonie van Leeuwenhoek* 1969, *35*, 325-343.
[112] Sydor, T., Schaffer, S., Boles, E., Considerable Increase in Resveratrol Production by Recombinant Industrial Yeast Strains with Use of Rich Medium. *Applied and environmental microbiology* 2010, *76*, 3361-3363.
[113] Duncan, K., Edwards, R. M., Coggins, J. R., The pentafunctional arom enzyme of *Saccharomyces cerevisiae* is a mosaic of monofunctional domains. *Biochemical Journal* 1987, *246*, 375–386.
[114] Duncan, K., Edwards, R. M., Coggins, J. R., The *Saccharomyces cerevisiae* ARO1 gene An example of the co-ordinate regulation of five enzymes on a single biosynthetic pathway. *FEBS Letters* 1988, *241*, 83-88.
[115] Patnaik, R., Liao, J. C., Engineering of *Escherichia coli* central metabolism for aromatic metabolite production with near theoretical yield. *Applied and environmental microbiology* 1994, *60*, 3903-3908.
[116] Lütke-Eversloh, T., Stephanopoulos, G., L-Tyrosine production by deregulated strains of *Escherichia coli*. *Applied microbiology and biotechnology* 2007, *75*, 103-110.
[117] Berry, A., Dodge, T. C., Pepsin, M., Weyler, W., Application of metabolic engineering to improve both the production and use of biotech indigo. *J Ind Microbiol Biotechnol* 2002, *28*, 127-133.
[118] Moxley, J. F., Jewett, M. C., Antoniewicz, M. R., Villas-Boas, S. G.*, et al.*, Linking high-resolution metabolic flux phenotypes and transcriptional regulation in yeast modulated by the global regulator Gcn4p. *Proceedings of the National Academy of Sciences of the United States of America* 2009, *106*, 6477-6482.
[119] Alper, H., Moxley, J., Nevoigt, E., Fink, G. R., Stephanopoulos, G., Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science (New York, N.Y.)* 2006, *314*, 1565-1568.

[120] Alper, H., Stephanopoulos, G., Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metabolic engineering* 2007, *9*, 258-267.
[121] Santos, C. N. S., Xiao, W., Stephanopoulos, G., Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 2012, *109*, 13538-13543.
[122] Mumberg, D., Müller, R., Funk, M., Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 1995, *156*, 119-122.
[123] Hegemann, J., Heick, S., Delete and Repeat: A Comprehensive Toolkit for Sequential Gene Knockout in the Budding Yeast *Saccharomyces cerevisiae*, in: Williams, J. A. (Ed.), *Strain Engineering*, Humana Press 2011, pp. 189-206.
[124] Leavitt, J. M., Alper, H. S., Advances and current limitations in transcript-level control of gene expression. *Current opinion in biotechnology* 2015, *34*, 98-104.
[125] Kushwaha, M., Salis, H. M., A portable expression resource for engineering cross-species genetic circuits and pathways. *Nat Commun* 2015, *6*.
[126] Farasat, I., Kushwaha, M., Collens, J., Easterbrook, M.*, et al.*, Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Molecular systems biology* 2014, *10*, 731-731.
[127] Biggs, B. W., De Paepe, B., Santos, C. N. S., De Mey, M., Kumaran Ajikumar, P., Multivariate modular metabolic engineering for pathway and strain optimization. *Current opinion in biotechnology* 2014, *29*, 156-162.
[128] Prindle, A., Selimkhanov, J., Li, H., Razinkov, I.*, et al.*, Rapid and tunable post-translational coupling of genetic circuits. *Nature* 2014, *508*, 387-391.
[129] Wang, B., Barahona, M., Buck, M., Engineering modular and tunable genetic amplifiers for scaling transcriptional signals in cascaded gene networks. *Nucleic acids research* 2014.
[130] Curran, K. a., Crook, N. C., Karim, A. S., Gupta, A.*, et al.*, Design of synthetic yeast promoters via tuning of nucleosome architecture. *Nature communications* 2014, *5*, 4002-4002.
[131] Curran, K. A., Karim, A. S., Gupta, A., Alper, H. S., Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metabolic engineering* 2013, *19*, 88-97.
[132] Levo, M., Segal, E., In pursuit of design principles of regulatory sequences. *Nat Rev Genet* 2014, *15*, 453-468.
[133] Levo, M., Zalckvar, E., Sharon, E., Dantas Machado, A. C.*, et al.*, Unraveling determinants of transcription factor binding outside the core binding site. *Genome Research* 2015.
[134] Teo, W. S., Chang, M. W., Bacterial XylRs and synthetic promoters function as genetically encoded xylose biosensors in *Saccharomyces cerevisiae*. *Biotechnology journal* 2015, *10*, 315-322.

[135] Zhang, M., Wang, F., Li, S., Wang, Y.*, et al.*, TALE: a tale of genome editing. *Progress in biophysics and molecular biology* 2014, *114*, 25-32.
[136] Ottoz, D. S. M., Rudolf, F., Stelling, J., Inducible, tightly regulated and growth condition-independent transcription factor in *Saccharomyces cerevisiae*. *Nucleic acids research* 2014, *42*, e130-e130.
[137] Dobrin, A., Saxena, P., Fussenegger, M., Synthetic biology: applying biological circuits beyond novel therapies. *Integrative Biology* 2016, *8*, 409-430.
[138] Solow, S. P., Sengbusch, J., Laird, M. W., Heterologous protein production from the inducible MET25 promoter in *Saccharomyces cerevisiae*. *Biotechnology progress* 2005, *21*, 617-620.
[139] Wimalarathna, R. N., Pan, P. Y., Shen, C.-H., Co-dependent recruitment of Ino80p and Snf2p is required for yeast CUP1 activation. *Biochemistry and Cell Biology* 2013, *92*, 69-75.
[140] Korber, P., Barbaric, S., The yeast PHO5 promoter: from single locus to systems biology of a paradigm for gene regulation through chromatin. *Nucleic acids research* 2014, *42*, 10888-10902.
[141] Weinhandl, K., Winkler, M., Glieder, A., Camattari, A., Carbon source dependent promoters in yeasts. *Microbial cell factories* 2014, *13*, 5-5.
[142] Li, Q., Zhao, X.-Q., Chang, A. K., Zhang, Q.-M., Bai, F.-W., Ethanol-induced yeast flocculation directed by the promoter of TPS1 encoding trehalose-6-phosphate synthase 1 for efficient ethanol production. *Metabolic engineering* 2012, *14*, 1-8.
[143] Williams, T. C., Averesch, N. J. H., Winter, G., Plan, M. R.*, et al.*, Quorum-sensing linked RNA interference for dynamic metabolic pathway control in *Saccharomyces cerevisiae*. *Metabolic engineering* 2015, *29*, 124-134.
[144] Lee, K., Hahn, J.-S., Interplay of Aro80 and GATA activators in regulation of genes for catabolism of aromatic amino acids in *Saccharomyces cerevisiae*. *Molecular Microbiology* 2013, *88*, 1120-1134.
[145] Lee, K., Sung, C., Kim, B.-G., Hahn, J.-S., Activation of Aro80 transcription factor by heat-induced aromatic amino acid influx in *Saccharomyces cerevisiae*. *Biochemical and biophysical research communications* 2013, *438*, 43-47.
[146] Kim, S., Lee, K., Bae, S.-J., Hahn, J.-S., Promoters inducible by aromatic amino acids and γ-aminobutyrate (GABA) for metabolic engineering applications in *Saccharomyces cerevisiae*. *Applied microbiology and biotechnology* 2015, *99*, 2705-2714.
[147] Williams, T. C., Nielsen, L. K., Vickers, C. E., Engineered Quorum Sensing Using Pheromone-Mediated Cell-to-Cell Communication in *Saccharomyces cerevisiae*. *ACS synthetic biology* 2013, *2*, 136-149.
[148] Teixeira, M. C., Monteiro, P. T., Guerreiro, J. F., Gonçalves, J. P.*, et al.*, The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in *Saccharomyces cerevisiae*. *Nucleic acids research* 2014, *42*, D161-D166.

[149] Redden, H., Alper, H. S., The development and characterization of synthetic minimal yeast promoters. *Nat Commun* 2015, *6*.
[150] Siuti, P., Yazbek, J., Lu, T. K., Synthetic circuits integrating logic and memory in living cells. *Nat Biotech* 2013, *31*, 448-452.
[151] Vargas-Tah, A., Gosset, G., Production of Cinnamic and p-Hydroxycinnamic Acids in Engineered Microbes. *Frontiers in Bioengineering and Biotechnology* 2015, *3*, 116.
[152] Rodriguez, A., Kildegaard, K. R., Li, M., Borodina, I., Nielsen, J., Establishment of a yeast platform strain for production of p-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. *Metabolic engineering* 2015, *31*, 181-188.
[153] Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A*., et al.*, De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. *Microbial cell factories* 2012, *11*, 155-155.
[154] Barrick, J. E., Lenski, R. E., Genome dynamics during experimental evolution. *Nat Rev Genet* 2013, *14*, 827-839.
[155] Reyes, L. H., Gomez, J. M., Kao, K. C., Improving carotenoids production in yeast via adaptive laboratory evolution. *Metabolic engineering* 2014, *21*, 26-33.
[156] Bonomo, J., Lynch, M. D., Warnecke, T., Price, J. V., Gill, R. T., Genome-scale analysis of anti-metabolite directed strain engineering. *Metabolic engineering* 2008, *10*, 109-120.
[157] Dmytruk, K. V., Yatsyshyn, V. Y., Sybirna, N. O., Fedorovych, D. V., Sibirny, A. A., Metabolic engineering and classic selection of the yeast *Candida famata* (*Candida flareri*) for construction of strains with enhanced riboflavin production. *Metabolic engineering* 2011, *13*, 82-88.
[158] Rogers, J. K., Taylor, N. D., Church, G. M., Biosensor-based engineering of biosynthetic pathways. *Current opinion in biotechnology* 2016, *42*, 84-91.
[159] Yang, J., Seo, S. W., Jang, S., Shin, S.-I*., et al.*, Synthetic RNA devices to expedite the evolution of metabolite-producing microbes. *Nat Commun* 2013, *4*, 1413.
[160] Fowden, L., Lewis, D., Tristram, H., Toxic Amino Acids: Their Action as Antimetabolites, *Advances in Enzymology and Related Areas of Molecular Biology*, John Wiley & Sons, Inc. 2006, pp. 89-163.
[161] Shetty, K., Crawford, D. L., Pometto, A. L., Production of l-Phenylalanine from Starch by Analog-Resistant Mutants of Bacillus polymyxa. *Applied and environmental microbiology* 1986, *52*, 637-643.
[162] Lee, D.-H., Feist, A. M., Barrett, C. L., Palsson, B. Ø., Cumulative Number of Cell Divisions as a Meaningful Timescale for Adaptive Laboratory Evolution of *Escherichia coli*. *PloS one* 2011, *6*, e26172.
[163] Winston, F., EMS and UV Mutagenesis in Yeast, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. 2001.
[164] Lütke-Eversloh, T., Stephanopoulos, G., A semi-quantitative high-throughput screening method for microbial L-tyrosine production in microtiter plates. *Journal of industrial microbiology & biotechnology* 2007, *34*, 807-811.

[165] Jensen, N. B., Strucko, T., Kildegaard, K. R., David, F.*, et al.*, EasyClone: method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. *FEMS yeast research* 2014, *14*, 238-248.
[166] DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X.*, et al.*, Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic acids research* 2013, *41*, 4336-4343.
[167] Richard, P., Viljanen, K., Penttilä, M., Overexpression of PAD1 and FDC1 results in significant cinnamic acid decarboxylase activity in *Saccharomyces cerevisiae*. *AMB Express* 2015, *5*, 1-5.
[168] Sambrook, J., in: Russell, D. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 2000.
[169] Boeke, J. D., Croute, F., Fink, G. R., A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Molecular and General Genetics MGG* 1984, *197*, 345-346.
[170] Hegeman, G. D., Synthesis of the Enzymes of the Mandelate Pathway by *Pseudomonas putida* I. Synthesis of Enzymes by the Wild Type. *Journal of Bacteriology* 1966, *91*, 1140-1154.
[171] Strøman, P., Reinert, W. R., Giles, N. H., Purification and characterization of 3-dehydroshikimate dehydratase, an enzyme in the inducible quinic acid catabolic pathway of *Neurospora crassa*. *Journal of Biological Chemistry* 1978, *253*, 4593-4598.
[172] Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I.*, et al.*, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature protocols* 2011, *6*, 1290-1307.
[173] Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C.*, et al.*, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth* 2009, *6*, 343-345.
[174] Gibson, D. G., Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic acids research* 2009, *37*, 6984-6990.
[175] Gietz, R. D. a. R. A. W., TRANSFORMATION OF YEAST BY THE Liac/SS CARRIER DNA/PEG METHOD. *Methods in Enzymology* 2002, 87-96.
[176] Teste, M.-A., Duquenne, M., François, J. M., Parrou, J.-L., Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Molecular Biology* 2009, *10*, 1-15.
[177] Lee, M. E., DeLoache, W. C., Cervantes, B., Dueber, J. E., A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS synthetic biology* 2015, *4*, 975-986.